高等学校经济管理类专业

实验教学系列教材

邢天才 主审



Jinrong Jiliangxue Shiyan

金融计量学实验

曲春青 主编

东北财经大学出版社
Dongbei University of Finance & Economics Press

★附配套光盘

邢天才　主审

高等学校经济管理类专业实验教学系列教材

Jinrong Jiliangxue Shiyan

金融计量学实验

曲春青　主编





定价：24.00元

高等学校经济管理类专业

实验教学系列教材

邢天才 主审



Jinrong Jiliangxue Shiyan

金融计量学实验

曲春青 主编

东北财经大学出版社
Dongbei University of Finance & Economics Press

图书在版编目（CIP）数据

金融计量学实验 / 曲春青主编. —大连：东北财经大学出版社，2008.8
（高等学校经济管理类专业实验教学系列教材）
ISBN 978 – 7 – 81122 – 442 – 9

Ⅰ.金… Ⅱ.曲… Ⅲ.金融–计量经济学–实验–高等学校–教材 Ⅳ.F830 – 33

中国版本图书馆 CIP 数据核字（2008）第 116945 号

东北财经大学出版社出版
（大连市黑石礁尖山街 217 号 邮政编码 116025）
总 编 室：（0411）84710523
营 销 部：（0411）84710711
网 址：http：//www.dufep.cn
读者信箱：dufep @ dufe.edu.cn
大连图腾彩色印刷有限公司印刷 东北财经大学出版社发行

幅面尺寸：170mm×240mm 字数：180 千字 印张：9 3/4
2008 年 8 月第 1 版 2008 年 8 月第 1 次印刷

责任编辑：李 彬 王 斌 责任校对：一 心
封面设计：冀贵收 版式设计：孟 鑫

ISBN 978 – 7 – 81122 – 442 – 9
定价：24.00 元

编审委员会

总 序

高等学校的教学原本就包括理论教学与实践教学两个部分。理论教学的任务主要是传授知识，教学的方法主要是教师讲授和在一定范围内的课堂讨论。实践教学的任务主要是培养能力，教学的方法主要是在教师的指导下由学生进行各种验证性、设计性实验和各种课程、专业及综合性的社会实践。高等教育由精英教育演进为大众教育后，在人才培养方面，为实现传授知识、培养能力和提高素质的统一，在注重理论教学的同时，开始赋予实践教学新的使命，并把实践教学推向了一个新的发展阶段。只有实践教学，才能验证知识，消化并巩固知识；只有实践教学，才能培养动手能力，形成专业素养；只有实践教学，才能启发创新思维，增强创新意识，提升创新能力。因此，在中国现阶段，无论是以培养应用型人才为主的教学型大学，还是以培养研究型人才为主的研究型大学，都把实践教学作为教学的重要组成部分。

从完整意义上说，实践教学包括实验教学和在教师指导下的社会实践两个部分。实验教学是以实验室为基地、在封闭或半封闭的条件下进行的实践教学，社会实践是在实习基地、在开放或有条件开放的条件下进行的实践教学。在人才培养方面，二者同等重要，不可或缺。然而，经济体制改革前，在“文科以社会为工厂”的指示指引下，理工科专业的实践教学主要采取了实验教学的模式，较少进行社会实践；文科专业的实践教学主要采取了社会实践的模式，较少进行实验教学。经济体制改革后，企、事业单位的经济地位和利益独立性发生了根本性的变化，政府的职能也发生了重要的转变，文科专业的社会实践遇到了“走出去”的困难。为了不降低人才培养质量，一些学校开始尝试性地建立实验室，组织实验教学，初步形成了实验教学与社会实践并行的实践教学体系。20 世纪 90 年代末后，随着招生规模的日益扩大，“走出去”实践的困难越来越大。为保证和提高人才培养质量，众多学校开始集中精力建设实验室，系统组织实验教学，基本形成了实验教学为主、社会实践为辅的实践教学体系。

经验表明，开展实验教学，至少需要具备两个方面的条件：一是实验室的建设；二是实验教学体系的确立。二者紧密联系，相互制约，相辅相成。实验教学体系的确立，必须以实验室的建设为前提。没有实验室的建设，实验教学体系的确立就会成为空中楼阁，流为空谈。同时，实验室的建设也必须服从实验教学体系建设的要求。离开了实验教学体系的要求，实验室的建设就会迷失方向，丧失标准。

抛开实验室建设，仅就实验教学体系建设而言，它大体上应包括实验课程的确

定和实验教学资源的开发与建设两个方面。一所高等学校，究竟确定哪些实验课程，开发和建设哪些实验教学资源，是由专业属性、人才培养的目标与规格及理论教学的体系与内容决定的。也就是说，实验教学体系的建设，必须同专业属性、人才培养的目标与规格及理论教学的体系与内容相符合。离开了专业属性、人才培养的目标与规格及理论教学的体系与内容，所建立的实验教学体系就没有灵魂、没有特色。

我校是以经济管理学科为主、培养高素质应用型人才的多学科性大学。从经济管理类专业的学科属性出发，按照培养高素质应用型人才的要求，参照理论教学的体系与内容，我们确立了“三层两级”的实验教学体系。这种实验教学体系，从课程设置来看，包括课程实验、专业实验、多专业协同实验三个层次，从实验内容来看，包括验证性实验和设计性实验两个级别。其中，课程实验在课程内进行，专业实验分专业单独设置，多专业协同实验由学校统一组织，验证性实验以验证知识、培养动手能力为主，设计性实验以启发创新思维、增强创新意识、提高创新能力为主。

为组织实施这种“三层两级”的实验教学，我们组织编写了这套“高等学校经济管理类专业实验教学系列教材”，具体包括《计算机应用基础实验》、《数据库基础实验》、《会计循环网络实验》、《统计学实验》、《证券、期货、外汇模拟实验》、《计量经济学实验》、《管理信息系统实验》等，共约30本。

为保证教材质量，我们专门成立了高等学校经济管理类专业实验教学系列教材编审委员会，建立了规范的编审制度。但由于缺少范式、经验不足、时间有限，教材中仍难免存在这样或那样的缺点或错误，诚恳地希望读者们指正。

马国维

2008年1月



前 言

金融计量学，是一门金融学与经济计量学相结合的交叉学科。在定量分析占据金融学研究主流的今天，金融计量学教学越来越受到各高等院校经济、金融专业师生的重视，同时金融计量学方法也在实际工作中得到越来越广泛的应用。

金融计量学是一门实践性很强的课程，不仅要求学生掌握金融计量学的基础理论知识，还要求学生学会如何将金融计量学理论知识应用到实际金融领域中去。因此，实验教学就成为金融计量学教学体系中不可或缺的一个重要组成部分。为了培养学生的实践操作能力，提高学生用定量分析的方法解决实际金融问题的水平，指导金融计量学的实验教学，我们编写了这本教材。

本书的编写分为实验篇和案例篇两个部分。根据金融计量学的主要教学内容，实验篇选择性地设计了七个实验，包括线性回归模型应用，简单外推模型应用，平滑技术和季节调整，随机时间序列的特征观察，单位根检验、协整与误差修正模型，ARMA 模型应用和条件异方差模型应用。每个实验包括实验目的与要求、实验准备知识、实验数据、实验内容、实验步骤、问题思考和实验总结等要点，对整个实验过程起到了重要的指导作用。其中的实验准备知识部分，是对实验所涉及的理论知识作一简要介绍，无法像单纯的理论教材那样详尽、具体，因此如果有需要还请读者参考金融计量学方面的理论教材。案例篇针对实验篇的七个实验，结合现实的金融计量问题，提供了七个案例，包括招商银行 A 股 β 系数估计，香港认可机构港元存款余额简单外推模型预测，我国金融机构财政存款余额平滑和季节调整，我国金融机构储蓄存款余额序列的特征观察，上证 A、B 股指数协整关系检验及误差修正模型，美元对欧元汇率 ARMA 模型应用和上证 A 股收益率条件异方差模型应用。案例篇的主要作用是辅助每个实验的教学以及为读者提供参考。

本书可以作为本科高年级或研究生阶段金融计量学实验的指导用书，也可以作为实际部门工作者自学金融计量学的参考用书。本书提供配套的资料光盘，内有实验和案例的数据，以及案例的 Eviews 工作文件，方便读者学习使用。

本书由曲春青编写初稿，邢天才主审，王振山、刘波对本书的结构和部分内容提出了建设性的意见，姜美华、王丽、屈超、张晓东、刘丹丹、刘佳参与了部分资料的收集整理工作。同时，本书的出版得到了东北财经大学实验教学中心和东北财经大学出版社的大力支持，在此，我们表示由衷的感谢。

本书在编写过程中吸收和借鉴了国内外专家、学者的一些观点和研究成果，在此向有关作者表示衷心感谢。由于编者的水平所限，书中定有不妥甚至错误之处，恳请读者批评指正。

编　者

2008 年 5 月

目　录



实验篇

实验一　线性回归模型应用

【实验目的与要求】

1. 准确掌握线性回归模型的基本形式和 OLS 估计的基本原理。
2. 熟练掌握线性回归模型的模型检验方法。
3. 熟练掌握线性回归模型的模型评价方法。
4. 学会利用线性回归模型对股票的 β 系数进行估计。
5. 在教师的指导下独立完成实验，得出正确的结果，并完成实验报告。

【实验准备知识】

线性回归模型，是以一定的经济理论为依据，定量刻画经济变量之间关系的一类模型。在相当长的时间里代表了计量经济学的主流，故又称为经典线性模型。该类模型一直是计量经济学课程的教学重点，也是大部分计量经济学教材的核心内容。因此，本书只对线性回归模型的基本内容作一简单回顾：

1. 线性回归模型的基本形式

线性回归模型根据解释变量的多少可分为一元线性回归模型和多元线性回归模型，而一元线性回归模型可以看作多元线性回归模型的一个特例，因此，我们这里只介绍多元线性回归模型。

假定被解释变量 Y 与 K 个解释变量 X_1，X_2，…，X_k 具有线性相关关系，多元线性回归模型的基本形式为：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i \qquad i=1, 2, 3, \cdots, n \tag{1.1}$$

用矩阵形式表达为：

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad (1.2)$$

令

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix},\quad X = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix},\quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix},\quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

可写成如下形式：

$$Y = X\beta + u \quad (1.3)$$

其中：Y 是被解释变量观测值的 n 维列向量；X 是所有自变量的 n 个样本点实际观测值构成的 $n\times(k+1)$ 矩阵；β 是 $k+1$ 维参数向量，由待估参数构成；u 是随机扰动向量。

2. 线性回归模型的基本假定

线性回归模型有如下基本假定：

（1）解释变量与随机扰动项不相关假定，即：

$$Cov(X_i, u_i) = 0 \quad (1.4)$$

但是，如果解释变量不是随机变量，这个假定将自动满足。

（2）随机扰动项零均值假定，即：

$$E(u_i) = 0 \quad (1.5)$$

（3）随机扰动项同方差假定，即：

$$Var(u_i) = \sigma^2 \quad (1.6)$$

（4）随机扰动项之间不相关假定，即：

$$Cov(u_i, u_j) = 0,\ i \neq j \quad (1.7)$$

（5）解释变量之间不存在线性相关关系假定，即：

$$Cov(X_i, X_j) = 0,\ i \neq j \quad (1.8)$$

（6）为了假设检验，假定随机误差项 u_i 服从均值为零，方差为 σ^2 的正态分布。即：

$$u_i \sim N(0, \sigma^2) \quad (1.9)$$

3. 线性回归模型的参数估计

线性回归模型参数估计最常用的方法是普通最小二乘法（OLS），得到的估计量称为最小二乘估计量。其基本思想是使被解释变量的观测值 Y_i 与模型估计值 $\hat{Y}_i$ 之差的平方和最小。参数估计值 $\hat{\beta}$ 为：

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (1.10)$$

在满足基本假设的前提下，该估计量是所有线性、无偏估计量中方差最小的，即 OLS 估计量是最优线性无偏估计量。

4. 线性回归模型的模型检验

线性回归模型的模型检验很多，常用的基本检验主要包括方程显著性检验（F 检验）、回归系数的显著性检验（t 检验）和残差检验等。

（1）方程显著性检验（F 检验）

方程的显著性检验（F 检验），主要是针对模型拟合样本的整体效果进行检验，即所有解释变量对被解释变量的总体解释力度。

检验的零假设为　　H_0：$\beta_0=\beta_1=\cdots=\beta_k=0$

检验的备择假设为　H_1：β_j（j=0，1，…，k），不全为零

在原假设成立条件下，统计量为：

$$F=\frac{SSR/k}{SSE/(n-k-1)}\sim F(k,n-k-1) \tag{1.11}$$

其中：$SSR=\sum(\hat{Y}_i-\bar{Y})^2$ 是回归平方和；$SSE=\sum(Y_i-\hat{Y}_i)^2$ 是残差平方和；被估参数个数为 $k+1$；n 是样本容量。

判别规则是：若 $F\le F_\alpha(k,n-k-1)$，接受 H_0；

若 $F>F_\alpha(k,n-k-1)$，拒绝 H_0。

从 Eviews 输出结果中看，如果 F 统计量的相伴概率小于 0.05，则在 5%的显著性水平上拒绝原假设，认为该回归的系数不全为零；如果 F 统计量的相伴概率大于 0.05，则在 5%的显著性水平上接受原假设，认为该回归的系数全为零。

（2）回归系数的显著性检验（t 检验）

回归系数的显著性检验（t 检验），主要是检验解释变量 X_j 对被解释变量 Y 是否具有显著影响，即 β_j 是否为零。如果 F 检验的结论是接受原假设，则不需再做 t 检验。如果 F 检验的结论是拒绝原假设，则进一步作 t 检验。

检验的零假设为　　H_0：$\beta_j=0$

检验的备择假设为　H_1：$\beta_j\neq0$，（j=0，1，…，k）

在原假设成立条件下，统计量

$$t=\frac{\hat{\beta}_j}{se(\hat{\beta}_j)}\sim t(n-k-1) \tag{1.12}$$

其中：$se(\hat{\beta}_j)$是估计量$\hat{\beta}_j$的标准差。

判别规则是：若$|t| \le t_{\alpha/2}(n-k-1)$，接受$H_0$；

若$|t| > t_{\alpha/2}(n-k-1)$，拒绝$H_0$。

从 Eviews 输出结果中看，如果 t 统计量的相伴概率小于 0.05，则在 5%的显著性水平上拒绝原假设，认为系数不为零，对应的解释变量应该包括在模型中；如果 t 统计量的相伴概率大于 0.05，则在 5%的显著性水平上接受原假设，认为系数为零，对应的解释变量不应包括在模型中。

（3）残差正态性检验

残差正态性检验可以利用 *JB*（Jarque-Bera）统计量进行。首先介绍样本标准差、偏度（skewness）和峰度（kurtosis）的定义。

对于序列（y_1，y_2，…，y_n），样本标准差$\hat{\sigma}$定义为：

$$\hat{\sigma}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{1.13}$$

偏度 S 定义为

$$S=\frac{1}{n}\sum_{i=1}^{n}(\frac{y_i-\bar{y}}{\hat{\sigma}})^3 \tag{1.14}$$

峰度 K 定义为

$$K=\frac{1}{n}\sum_{i=1}^{n}(\frac{y_i-\bar{y}}{\hat{\sigma}})^4 \tag{1.15}$$

其中：y_i是观测值；$\bar{y}$是样本平均数；n 是样本容量。由公式（1.14）知，若分布是以$\bar{y}$对称的，则偏度 S 为零。所以若 y_t 服从正态分布，则偏度 S 为零；若分布是右偏倚的，则偏度 $S>0$；若分布是左偏倚的，则偏度 $S<0$。正态分布的峰度值为 3。如果一个分布的两侧尾部比正态分布的两侧尾部“胖”，则该分布的峰度 $K>3$，反之则 $K<3$。

JB 统计量定义为：

$$JB=\frac{n-m}{6}[S^2+\frac{1}{4}(K-3)^2]\sim\chi^2(2) \tag{1.16}$$

其中：m 是产生样本序列使用到的估计系数的个数；S 是偏度；K 是峰度。对于直接得到的观测时间序列，m=0。对于残差序列，m 等于原回归模型中解释变量个数。可以通过计算 *JB* 统计量的值，来检验一个分布（比如残差的分布）是否为正态。

判别规则是：若 $JB \le \chi_{\alpha}^2(2)$，该分布为正态分布；

若 $JB > \chi_{\alpha}^2(2)$，该分布不是正态分布。

从 Eviews 输出结果中看，如果 *JB* 统计量的相伴概率大于 0.05，则在 5%的显著

性水平上认为该分布为正态分布；如果 JB 统计量的相伴概率小于 0.05，则在 5%的显著性水平上认为该分布不是正态分布。

（4）残差序列自相关的 DW 检验

杜宾—瓦森检验，简称 DW 检验，是检验残差序列自相关最常用的解析方法，但该检验只适用于检验一阶自相关。同时，在回归模型的解释变量中不能含有被解释变量的滞后项，否则 DW 检验将失效。DW 统计量为：

$$DW=\sum_{i=2}^{n}(\hat{u}_i-\hat{u}_{i-1})^2/\sum_{i=1}^{n}\hat{u}_i^{\,2} \tag{1.17}$$

其中：n 是样本容量；$\hat{u}_i$ 是残差。DW 的值域为[0，4]，在给定的显著性水平下，可以查表得到 DW 的下界值 DW_L 和上界值 DW_U，从而确定判断范围。

判别规则是：

若 $0<DW<DW_L$，$\hat{u}_i$ 存在一阶正自相关，DW 越接近 0，正自相关性越强。

若 $4-DW_L<DW<4$，$\hat{u}_i$ 存在一阶负自相关，DW 越接近 0，负自相关性越强。

若 $DW_U<DW<4-DW_U$，$\hat{u}_i$ 不存在自相关，DW 越接近 2，判断无自相关性把握越大。

若 $DW_L<DW<DW_U$ 或者 $4-DW_U<DW<4-DW_L$，不能确定 $\hat{u}_i$ 是否存在自相关。

（5）残差序列自相关的 LM 检验

LM 检验的特点是既可检验一阶自相关，也可检验高阶自相关。同时，回归模型的解释变量中含有被解释变量的滞后项，LM 检验仍然有效。

LM 检验是通过建立一个辅助回归式，构造 LM 统计量进行自相关检验的。考虑随机扰动项 u_i 为 p 阶自回归形式

$$u_i=\rho_1 u_{i-1}+\rho_2 u_{i-2}+\cdots+\rho_p u_{i-p}+v_i \tag{1.18}$$

其中：v_i 是随机项，符合各种假定条件。

用多元线性回归估计式得到的残差 $\hat{u}_t$ 建立辅助回归式为：

$$\hat{u}_i=\rho_1\hat{u}_{i-1}+\cdots+\rho_p\hat{u}_{i-p}+\beta_0+\beta_1 X_{1i}+\cdots+\beta_k X_{ki}+v_i \tag{1.19}$$

估计上式，并计算样本决定系数 R^2。

LM 检验的零假设和备择假设为：

H_0 为 $\rho_1=\rho_2=\cdots=\rho_p=0$（$u_t$ 不存在 p 阶自相关）；H_1 为 ρ_j（j=0，1，…，p），

不全为零（u_t存在p阶自相关）。

构造LM统计量为：

$$LM = nR^2 \sim \chi^2(p) \tag{1.20}$$

其中：n是原多元线性回归估计式的样本容量；p是（1.18）式中自回归阶数。

判别规则是：若$LM \le \chi^2_\alpha(p)$，接受H_0；

若$LM > \chi^2_\alpha(p)$，拒绝H_0。

从Eviews输出结果中看，如果LM统计量的相伴概率小于0.05，则在5%的显著性水平上拒绝原假设，认为原回归式中的u_i存在p阶自相关；如果LM统计量的相伴概率大于0.05，则在5%的显著性水平上接受原假设，认为原回归式中的u_i不存在p阶自相关。

（6）残差序列异方差的White检验

White检验由H. White于1980年提出，它是把残差平方$\hat{u}_i^2$作为被解释变量，把原回归方程的解释变量和解释变量的平方项作为新的解释变量建立线性回归模型（还可以加上任意两个解释变量的交叉积项）。通过这个辅助回归式构造χ^2统计量进行异方差检验。以二元回归模型为例：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i \tag{1.21}$$

首先对上式进行OLS估计，求出残差$\hat{u}_i$，然后做如下辅助回归式：

$$\hat{u}_i^2 = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_3 X_{1i}^2 + \alpha_4 X_{2i}^2 + \alpha_5 X_{1i} X_{2i} + v_i \tag{1.22}$$

即用$\hat{u}_i^2$对原回归式中的各解释变量、解释变量的平方项、交叉积项建立线性回归模型。估计上式，并求出样本决定系数R^2。

White检验的零假设和备择假设为：

H_0为原回归式中的u_i不存在异方差；H_1为原回归式中的u_i存在异方差。

在零假设成立的条件下统计量为：

$$nR^2 \sim \chi^2(k) \tag{1.23}$$

其中：n是样本容量；自由度k是辅助回归式中解释变量个数（不计算常数项），这里k应取5。

判别规则是：若$nR^2 \le \chi^2_\alpha(k)$，接受$H_0$；

若 $nR^2 > \chi_\alpha^2(k)$，拒绝 H_0。

从 Eviews 输出结果中看，如果 nR^2 的相伴概率小于 0.05，则在 5%的显著性水平上拒绝原假设，认为原回归式中的 u_i 存在异方差；如果 nR^2 的相伴概率大于 0.05，则在 5%的显著性水平上接受原假设，认为原回归式中的 u_i 不存在异方差。

5. 线性回归模型的模型评价

模型评价的主要目的是评价一个模型的优劣，或在几个不同模型之间进行比较和选择。评价的指标主要有三种。

（1）样本决定系数 R^2 和调整后的 R^2

通常我们可以将总离差平方和（SST）分解为回归平方和（SSR）与残差平方和（SSE）之和的形式，即：

$$\sum_{i=1}^{n}(Y_i-\bar{Y})^2=\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2+\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2 \tag{1.24}$$

或 $SST = SSR + SSE$。

对于一个拟合得好的模型，总离差平方和与回归平方和应该较为接近。因此，可以选择二者接近程度作为评价模型拟合优度的标准。其定义为：

$$R^2=\frac{SSR}{SST}=\frac{\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2} \tag{1.25}$$

其中：R^2 是样本决定系数，它表示了总离差平方和中由回归方程可以解释的部分所占的比例。

显然 R^2 是一个介于 0 到 1 之间的数，越接近 1 说明拟合效果越好。但是，R^2 通常会随着解释变量的个数增加而增大，然而新增变量会降低自由度，因此单凭 R^2 增大而认为拟合效果越来越好是不明智的。为了克服这个缺点，引进调整后的 R^2，记为 $\bar{R}^2$。其定义为：

$$\bar{R}^2=1-\frac{n-1}{n-k}(1-R^2) \tag{1.26}$$

其中：n 是样本容量；k 是被估计参数个数。该指标综合了精度和变量数两个因素，兼顾了精确性和简洁性，用来度量拟合优度比 R^2 更有效。$\bar{R}^2$ 也是一个介于 0 到 1 之间的数，越接近 1 说明拟合效果越好。

（2）对数似然值

对数似然值（log-likelihood，简记为 L），是基于极大似然估计方法得到的统计量。它的形式为：

$$L = -\frac{n}{2}[1 + \log(2\pi) + \log(\frac{\hat{u}'\hat{u}}{n})] \tag{1.27}$$

其中：n 是样本容量；$\hat{u}$ 是残差向量；$\hat{u}'\hat{u}$ 是残差平方和。可见，残差平方和越小，L 越大，模型越精确，但是解释变量的个数越多，残差也越小，因此 L 和样本决定系数 R^2 有同样的缺陷。

（3）赤池信息准则（AIC 准则）和施瓦茨准则（SC 准则）

AIC 准则和 SC 准则，可以看做对对数似然值（L）的改进，在对几个模型进行比较选择，尤其是在确定一个滞后分布的长度时，被经常使用。

赤池信息准则（AIC 准则）定义为：

$$AIC = -\frac{2L}{n} + \frac{2k}{n} \tag{1.28}$$

施瓦茨准则（SC 准则）定义为：

$$SC = -\frac{2L}{n} + \frac{k\log n}{n} \tag{1.29}$$

其中：L 是对数似然值；k 是被估计参数个数；n 是样本容量。

容易看出，AIC 和 SC 的值随着 L 的增大而变小，随着 k 的增大而变大。而 k 的增大又会使 L 变大，因此，随着 k 的变化，AIC 和 SC 有极小值存在。使用 AIC 准则和 SC 准则的方法是通过连续增加解释变量个数直到 AIC 和 SC 取得极小值，从而确定最优的 k 值。如果说 L 代表了模型的精确性，k 代表了模型的简洁性，那么 AIC 准则和 SC 准则可以说兼顾了精确性和简洁性。

6. 不满足基本假定的问题及对策[①]

在满足基本假设的前提下，OLS 估计量是最优线性无偏估计量。但是在实际的计量经济问题中，完全满足这些基本假设的情况并不多见。如果违背了其中某一项或几项基本假设，那么就需要使用新的方法来估计模型。

（1）多重共线性问题

多重共线性是指在多元回归中，两个或多个解释变量之间高度相关甚至是完全线性相关。

①多重共线性的后果。如果存在多重共线性问题，将会产生：回归参数估计的标准差变大；置信区间变宽；t 值不显著；R^2 很高，但是一些重要解释变量参数的 t 检验通不过；OLS 估计对数据的微小变化非常敏感，即稳定性降低；回归系数符号有误；难以衡量各解释变量对回归平方和或者 R^2 的贡献等。

②多重共线性的诊断。一般情况下，我们可以通过一些经验方法诊断多重共线性。常用的方法有：R^2 较高，但是 t 值显著的不多；解释变量之间样本相关系数很高；回归系数的符号与经济理论相反等。

③多重共线性的补救措施。常用的补救措施有：从模型中删除不重要的解释变

---

① 由于此部分涉及的问题比较繁杂，本书只做简单介绍，有兴趣的读者可参考有关的经济计量学教材。

量；获取额外的数据或者新的样本；重新设定模型；变换模型中的变量等。

（2）异方差问题

如果回归模型中的随机扰动项不满足同方差假定，即：

$$Var(u_i) \neq Var(u_j)\ ,\ i \neq j \tag{1.30}$$

则称 u_i 存在异方差性。

①异方差性的后果。如果存在异方差问题，将会产生：OLS 估计虽然是线性无偏的，但是不再具有最小方差性，即不是有效的；OLS 估计量的方差是有偏的；基于 t 分布和 F 分布之上的置信区间和假设检验不再有效等。

②异方差性的诊断。可以利用前面介绍的 White 检验对异方差问题进行诊断。

③异方差问题的补救措施。如果模型被诊断出异方差问题，可采用加权最小二乘法和广义最小二乘法对模型进行估计。

（3）自相关问题

如果回归模型中的随机扰动项不满足不相关的假定，即：

$$Cov(u_i, u_j) \neq 0\ ,\ i \neq j \tag{1.31}$$

则称 u_i 存在自相关现象，也称序列相关。

①自相关的后果。如果存在自相关问题，将会产生：OLS 估计虽然是线性无偏的，但不是有效的；OLS 估计量的方差是有偏的；t 检验和 F 检验不再有效；计算得到的误差方差是真实方差的有偏估计量；R^2 无效；通常计算的预测的方差和标准差也是无效的等。

②自相关的诊断。自相关的诊断可以用前面介绍的 DW 检验和 LM 检验。

③自相关问题的补救措施。如果模型被诊断出自相关问题，可采用广义最小二乘法及两步最小二乘法进行估计，还可以采用 ARIMA 模型进行修正。

7. 股票 β 系数估计的线性回归模型应用

资本资产定价模型（CAPM）认为，总体风险等于系统风险加上非系统风险，而非系统风险可以通过一个充分分散化的投资组合消除，剩下的就只有系统风险。那么某个证券或证券组合的预测收益就是由它所含有的系统风险唯一确定的，可表示为：

$$E(r_i) = r_f + [E(r_m) - r_f]\beta_i \tag{1.32}$$

其中：$E(r_i)$ 表示证券 i 的期望收益率；r_f 表示无风险收益率；$E(r_m)$ 表示市场组合的期望收益率。这里 $\beta_i = \text{cov}(r_i, r_m)/\text{var}(r_m) = \sigma_{im}/\sigma_m^2$，称为 β 系数，表示证券 i 的风险相对于市场风险的比率，反映证券 i 的收益率对市场收益率的敏感性，即该证券系统风险的大小。

在实际计量中，我们经常通过某证券或证券组合收益率和市场组合收益率的历史数据，利用线性回归模型来估计该证券或证券组合的 β 系数。比如用 r_i 表示 i 股票

的收益率，r_m 表示市场组合的收益率（通常可以用市场指数收益率替代），可以利用单指数模型建立一元线性回归模型：

$$r_{it} = a + br_{mt} + u_t \tag{1.33}$$

易知，系数 b 的估计值 $\hat{b}$ 即可看做该股票 β 系数 β_i 的估计值。

进一步，由单指模型可以得到：

$$E(r_i) = a + \beta_i E(r_m) \tag{1.34}$$

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_u^2 \tag{1.35}$$

即 i 股票的总体风险=系统风险+非系统风险。因此，i 股票的系统风险占总风险的比例为 $\frac{\beta_i^2 \sigma_m^2}{\sigma_i^2}$。同时，单指数模型的样本决定系数 R^2 为：

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{t=1}^{T}(\hat{r}_{it} - \overline{r}_i)^2}{\sum_{t=1}^{T}(r_{it} - \overline{r}_i)^2} = \frac{\hat{b}^2 \sum_{t=1}^{T}(r_{mt} - \overline{r}_m)^2}{\sum_{t=1}^{T}(r_{it} - \overline{r}_i)^2} = \frac{\hat{\beta}_i^2 \sigma_m^2}{\sigma_i^2} \tag{1.36}$$

可见，单指数模型中，样本决定系数 R^2 除了可以评价模型的拟合效果，还可以度量 i 股票的系统风险占总风险的比例。

【实验数据】

我国上海证券交易所上证指数（代码 000001）和宝钢股份（代码 600019）2002 年 4 月到 2008 年 2 月的月收盘价数据（参见资料光盘：数据集/线性回归模型应用数据/上证指数月数据.xls 和宝钢股份月数据.xls），数据来源为大智慧软件（下载并整理）。

【实验内容】

β 系数，是度量某证券或证券组合系统风险的重要指标，在实际的投资策略选择中起着非常重要的作用。β 系数估计的通用方法是利用某证券或证券组合收益率

和市场组合收益率的历史数据，通过线性回归模型来进行估计。本次实验，同学们可以根据我国上海证券交易所上证指数和宝钢股份 2002 年 4 月到 2008 年 2 月的月收盘价数据，建立一元线性回归模型，对宝钢股份的 β 系数进行估计。

【实验步骤】

1. 根据数据频率和时间范围，创建 Eviews 工作文件（Workfile）。

2. 录入数据，并对序列进行初步分析。分别绘制上证指数和宝钢股份月收盘价序列的折线图，初步分析序列的基本趋势和波动特征。

3. 相关性分析。建立上证指数和宝钢股份的百分比收益率序列，并求出两个序列的相关系数，绘制二维散点图，分析两个序列的相关性。

4. 建立单指数模型，并利用 OLS 方法进行参数估计。

5. 对模型进行检验。利用 t 统计量、DW 统计量、残差正态性检验、自相关的 LM 检验、异方差的 White 检验等方法对模型进行检验。

6. 对模型进行评价。利用样本决定系数 R^2 和调整后的 R^2 等指标结合模型的检验结果对模型进行评价。

7. 对模型的经济意义进行解释。对参数估计值和样本决定系数 R^2 的经济意义进行解释。

8. 综合上述实验步骤得出的结果，得出最终结论。总结实验过程中的问题以及得到的经验教训，完成实验报告。

【问题思考】

1. 线性回归模型的基本形式是什么样的？有哪些基本假定？
2. 普通最小二乘法（OLS）的基本原理是什么？
3. 线性回归模型的模型检验和模型评价的方法、指标有哪些？如何使用？
4. 如何估计某证券或证券组合的 β 系数？
5. 单指数模型的样本决定系数 R^2 有何经济含义？

【实验总结】

通过本实验，加深对线性回归模型基本形式和OLS估计基本原理的理解和掌握。熟练掌握线性回归模型的模型检验和模型评价方法。学会利用线性回归模型对股票的β系数进行估计。写出实验报告。

实验二　简单外推模型应用

【实验目的与要求】

1. 准确掌握简单外推模型的各种形式和基本原理。
2. 熟练掌握运用 Eviews 软件建立简单外推模型。
3. 学会利用多种方法对多个模型进行比较分析。
4. 熟练掌握运用简单外推模型对样本序列进行外推预测。
5. 在教师的指导下独立完成实验，得出正确的结果，并完成实验报告。

【实验准备知识】

1. 使用简单外推模型的理由

简单外推模型虽然不像结构模型那样具有较高的预测精度，但由于其方法简单、便于操作，在许多情况下也得到了广泛应用，比如：需要在很短时间对大规模的时间序列进行预测，时间和资源上不允许使用正规的模型技术；当有理由认为一个特定的时间序列有一个简单的趋势，也不必使用较复杂的模型。可以说，简单外推模型作为一种简单、便宜和可接受的预测手段是非常有用的，希望同学们能够熟练掌握和运用。

2. 简单外推模型简介

基于时间序列的过去数据来进行预测的简单模型，通常被称为确定性模型，它们不反映时间序列的随机性质。这类模型所涉及的外推技术多年来一直是经济、金融和商业预测的标准工具。

当我们确定时间序列 y_t 的一个基本特征是长期增长模式，同时确信它有一个确定的趋势，并将持续下去，我们就可以建立一个反映这一趋势的简单模型，并预测 y_t。下面简要介绍几个简单外推模型。

（1）线性趋势模型

线性趋势模型，即直线模型，基本形式为：

$$y_t = c_1 + c_2 t \tag{2.1}$$

其中：t 是时间；y_t 是 y 在 t 时期的值。

如果我们认为序列 y_t 在每个时期是以同样的增长量增长，就可以通过该模型来预测 y_t。通常 t 在基期（第一个观察值）取 1，且每增加一个时期增加 1。例如，回归得到：

$$y_t = 120 + 30t \tag{2.2}$$

即说明该序列每一期比前期增长 30 个单位，从而可以通过在 y_t 的水平上增加 30 个单位作为 y_{t+1} 期值的预测。

（2）指数增长模型

指数增长模型，基本形式为：

$$y_t = f(t) = Ae^{rt} \tag{2.3}$$

如果我们认为序列 y_t 在每个时期是以同样的百分比增长速度增长，就可以通过该模型来预测 y_t，由（2.3）式可知 $y_{t+1} / y_t = e^r$，所以该序列的发展速度为 e^r，增长速度为 $e^r - 1$。参数 A 和 r 的估计方法是将方程（2.3）两边取对数，并拟合对数线性方程：

$$\log y_t = c_1 + c_2 t \tag{2.4}$$

其中：$c_1 = \log A$；$c_2 = r$。

（3）自回归趋势模型

自回归趋势模型，即当期值只受前期值的影响，序列是其滞后项的函数。基本形式为：

$$y_t = c_1 + c_2 y_{t-1} \tag{2.5}$$

当c_1=0 时，c_2代表序列y_t的发展速度。当c_1不等于零，c_2=1 时，则序列y_t在每个时期都有相同的绝对增长量。

（4）对数自回归趋势模型

对数自回归趋势模型实质上是自回归趋势模型的变形，基本形式为：

$$\log y_t = c_1 + c_2 \log y_{t-1} \tag{2.6}$$

当c_1=0 时，c_2代表序列y_t的复合增长速度。

下面介绍两个相对复杂的简单的外推模型。

（5）二次趋势模型

二次趋势模型，是线性趋势模型的推广，即加了一个t^2项，基本形式为：

$$y_t = c_1 + c_2 t + c_3 t^2 \tag{2.7}$$

当c_2>0、c_3>0，则y_t总是不断增长，且加速。

当c_2<0、c_3>0，则y_t先递减然后递增。

当c_2<0、c_3<0，则y_t总是递减。

（6）逻辑曲线模型

逻辑曲线模型，代表一条 S 形曲线，基本形式为：

$$y_t = \frac{1}{k + ab^t} \quad b > 0 \tag{2.8}$$

方程关于参数 k，a，b 是非线性的，要用非线性估计方法估计。最典型应用逻辑曲线模型的例子，就是 k>0，a>0，0<b<1 用于反映在未来某天市场将饱和的产品的销售量，如图 2—1 所示。

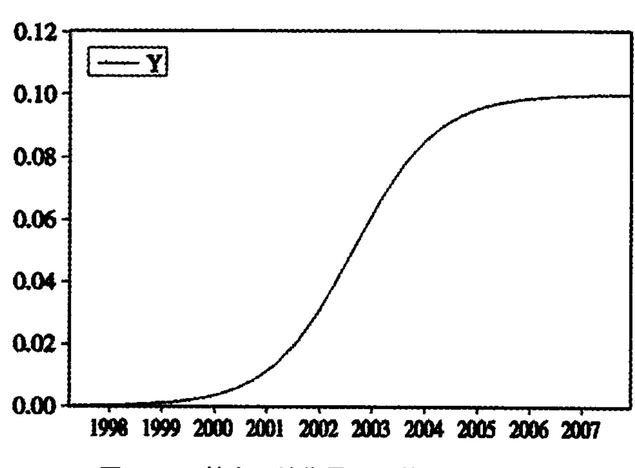


图 2—1 某产品销售量图（单位：亿件）

该销售量的趋势是早期为缓慢上升，中期为加速上升，后期为减速上升，并渐渐平缓接近一个极限值。

【实验数据】

我国金融机构企业存款余额月度数据，样本范围2000年1月到2007年12月（参见资料光盘：数据集/简单外推模型数据/金融机构企业存款余额月度数据.xls）。数据来源为中经网统计数据库。

【实验内容】

企业存款作为金融机构一项重要的负债业务，一直以来受到各金融机构的普遍重视。企业存款余额不仅能够反映企业的财务状况、经营能力、发展态势，更能够为金融机构制定发展规划提供重要的参考依据。本次实验，同学们可以根据我国金融机构企业存款余额2000年1月到2007年12月数据，分别运用线性趋势模型、指数增长模型、自回归趋势模型、对数自回归趋势模型、二次趋势模型进行拟合，并比较模型的优劣，最后选择最优模型对企业存款2008年1月余额进行外推预测。

【实验步骤】

1. 根据数据频率和时间范围，创建Eviews工作文件（Workfile）。

2. 录入数据，并对序列进行初步分析。绘制金融机构企业存款余额序列的折线图，分析序列的基本趋势。

3. 建立简单外推模型。分别建立线性趋势模型、指数增长模型、自回归趋势模型、对数自回归趋势模型、二次趋势模型对原序列进行回归拟合，对每个模型进行简单评价和检验。同时，绘制真实值、拟合值与残差图。

4. 对几个模型进行预测评价。分别用线性趋势模型、指数增长模型、自回归趋

势模型、对数自回归趋势模型、二次趋势模型进行静态预测，比较预测精度。观察2007年12月的真实值与预测值，计算预测误差，并比较。

5. 模型比较。(1) 图形比较。将上述几个模型的拟合值与原序列放在一个组里，以线性图形形式打开，观察哪个模型拟合得最好。将上述几个模型的预测值与原序列放在一个组里，以线性图形形式打开，观察哪个模型预测得最好。(2) 指标数据比较。将上述几个模型得到的拟合优度 R^2，有无一阶自相关，平均绝对百分比误差（MAPE），2007年12月数据的预测误差 η 汇总到一起，综合比较模型优劣。

6. 利用上一步得到的最优模型，对企业存款2008年1月余额进行外推预测。

7. 综合上述实验步骤得出的结果，得出最终结论。总结实验过程中的问题以及得到的经验教训，完成实验报告。

【问题思考】

1. 为什么要利用简单外推模型对序列进行外推预测？

2. 如果序列没有一个明显的趋势，是否还能使用简单外推模型？为什么？

3. 能否运用简单外推模型进行长期预测？比如预测2008年12月的企业存款余额。预测精度会如何？

【实验总结】

通过本实验，加深对几种简单外推模型形式和基本原理的理解和掌握。熟练运用多种手段对多个模型进行评价和比较。掌握运用简单外推模型进行外推预测的基本方法。写出实验报告。

实验三　平滑技术和季节调整

【实验目的与要求】

1. 准确掌握平滑技术和季节调整的各种形式和基本原理。
2. 熟练掌握运用 Eviews 软件对样本序列进行移动平均平滑。
3. 学会利用加法模型和乘法模型对样本序列进行季节调整。
4. 熟练掌握运用指数平滑方法对样本序列进行外推预测。
5. 在教师的指导下独立完成实验，得出正确的结果，并完成实验报告。

【实验准备知识】

平滑技术，是消除或至少减少时间序列短期波动的一个手段。这样做不仅可以使我们容易识别序列的趋势和周期变动类型，而且平滑后进行直觉的分析也会简单。季节调整是平滑的一种特殊形式，它消除时间序列季节波动的影响，更好地反映时间序列的运动规律。

1. 简单移动平均方法

对于时间序列 y_t，n 期简单移动平均公式为：

$$\tilde{y}_t = \frac{1}{n}(y_t + y_{t-1} + \cdots + y_{t-n+1}) \tag{3.1}$$

其中：$\tilde{y}_t$ 为简单移动平均平滑后的序列；n 为简单移动平均的期数。

n 期简单移动平均应用非常广泛，比如在股票投资分析中，经常看到的 5 日均

线、10 日均线、30 日均线，都是为了更好地看清股价的走势，而对股价序列进行简单移动平均后得到的。当然，n 越大，平滑的程度越高，越能体现长期的趋势。

2. 季节调整

季节变动是指以一年为一个周期的变化。时间序列的季度、月度观测值常常出现季度或月度的循环变动。这种变动的影响因素主要是四季更迭，还有人文或制度等方面的因素。季节变动往往会掩盖经济发展的客观规律，妨碍我们对某些问题的认识。因此，通常在利用季度或月度数据进行分析之前，我们需要对时间序列进行季节调整。

传统的时间序列分析把时间序列的波动归结为四大因素：趋势变动（L）、季节变动（S）、循环变动（C）和不规则变动（I）。循环变动指周期为数年的变动，通常指经济周期。不规则变动即随机变动。四种变动与原序列（Y）的关系被概括成两种模型：

乘法模型 $Y=L\times S\times C\times I$ (3.2)

加法模型 $Y=L+S+C+I$ (3.3)

其中：乘法模型适用于 L、S、C 相关的情形，比如，季节变动的幅度随趋势上升而增加；加法模型则适用于 L、S、C 相互独立的情形。

季节调整的基本思路就是将季节变动 S（季节因子，又称季节指数）从序列中去除。以乘法模型为例，首先，我们剔除长期趋势和循环变动的结合项 $L\times C$，我们可以用移动平均 $\tilde{y}_t$ 作为 $L\times C$ 的估计值，因为我们可大致认为 $\tilde{y}_t$ 已无季节和不规则波动。这里的 $\tilde{y}_t$ 是中心化的移动平均，即：

$$\tilde{y}_t=\begin{cases}(0.5Y_{t+6}+\cdots+Y_t+\cdots+0.5Y_{t-6})/12, & \text{月度数据}\\(0.5Y_{t+2}+Y_{t+1}+Y_t+Y_{t-1}+0.5Y_{t-2})/4, & \text{季度数据}\end{cases} \tag{3.4}$$

然后，我们用原序列除以 $L\times C$ 的估计值 $\tilde{y}_t$ 就得到季节和不规则变动的结合项 $S\times I$ 的一个估计：

$$\frac{L\times S\times C\times I}{L\times C}=S\times I=\frac{y_t}{\tilde{y}_t}=z_t \tag{3.5}$$

下一步尽可能从 z_t 中彻底消除 I，得到季节因子 S。由于对同一月份或季度的季节和不规则变动的结合项进行平均将大体上消除不规则变动，于是我们对 $S\times I$ 同一月份的数据进行平均，得到平均值 $\tilde{z}_j$，就可以作为季节指数的估计值。这里我们需要将季节因子标准化，方法为：

$$S_j=\tilde{z}_j/\sqrt[k]{\tilde{z}_1\bullet\tilde{z}_2\bullet\cdots\bullet\tilde{z}_k} \tag{3.6}$$

S_j 称为标准化的季节因子，对于季度数据 j=1，2，3，4，k=4；对于月度数据 j=1，2，…，12，k=12。调整后季节因子的乘积等于 1 。

最后我们来消除季节变动。从每个原序列数值中除以对应的季节因子 S_j，消除季节变动成分后，剩下其他三部分。

季节调整后的序列即：$y_t^a = y_t / S_j$　　(3.7)

对于加法模型，只要做一些对应的变化即可。将（3.5）中的除法变成减法，即 $y_t - \tilde{y}_t = z_t$；将(3.6)变为 $S_j = \tilde{z}_j - \frac{1}{k}\sum_{i=1}^{k}\tilde{z}_i$，使得调整后季节因子的和等于 0；将(3.7)中的除法变成减法，即：$y_t^a = y_t - S_j$。

Eviews3.1 软件中还提供了 Census X11 方法，更高版本中还提供了 Census X12 方法，这里不作详细介绍，有兴趣的读者可参看 Eviews 操作说明。

3. 指数平滑方法

（1）一次指数平滑（single exponential smoothing）

从公式（2.1）可以看到，简单移动平均将每期的权重赋予相同的值。然而我们通常认为 y_t 的近期值比早期的值更重要才合乎情理，即近期值应有更大的权重。因此，我们引入指数加权移动平均模型，其基本形式为：

$$\tilde{y}_t = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \cdots \tag{3.8}$$

其中：$\tilde{y}_t$ 为平滑之后的序列。将 $\tilde{y}_t$ 换为 $\tilde{y}_{t-1}$ 方程左右乘以 $(1-\alpha)$ 得到：

$$(1-\alpha)\tilde{y}_{t-1} = \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \cdots \tag{3.9}$$

两式相减，得到计算 $\tilde{y}_t$ 的迭代公式：

$$\tilde{y}_t = \alpha y_t + (1-\alpha)\tilde{y}_{t-1} \tag{3.10}$$

这里 $0 \le \alpha \le 1$，叫做平滑系数，又叫衰减因子。可以看到，α 越接近 1，现值越重要，即对应的权重越大。因此 α 越小，时间序列的平滑程度就越高。

上面就是我们通常所说的一次指数平滑，是实际值序列的加权平均，适用于比较平稳的序列，即序列值在一个常数均值上下随机波动，无趋势及季节要素的情况。由于权数成指数衰减，越早的数据被赋予越小的权重，因此预测值主要倚重近期样本数据，远期数据对它影响较小。同时，与实际序列的变化相比有滞后现象。

一次指数平滑的预测公式为：

$$\tilde{y}_{T+k} = \tilde{y}_T \quad \text{对于所有 } k \ge 1 \tag{3.11}$$

其中：T 是时间序列的最末期。

（2）二次指数平滑（double exponential smoothing）

如果我们希望被平滑的光滑程度较高，但又不对历史数据加权过重，即使α较小可能也达不到要求。此时可使用二次指数平滑，即对一次平滑好的序列$\tilde{y}_t$再平滑一次。其计算公式为：

$$\tilde{y}_t = \alpha y_t + (1-\alpha)\tilde{y}_{t-1}$$

$$\tilde{\tilde{y}}_t = \alpha \tilde{y}_t + (1-\alpha)\tilde{\tilde{y}}_{t-1} \tag{3.12}$$

其中：$\tilde{y}_t$是一次指数平滑序列；$\tilde{\tilde{y}}_t$是二次指数平滑序列；α是平滑系数，$0 \le \alpha \le 1$。

二次指数平滑的预测公式为：

$$\tilde{y}_{T+k} = a_T + b_T k \quad 对于所有\ k \ge 1 \tag{3.13}$$

其中：$a_T = 2\tilde{y}_T - \tilde{\tilde{y}}_T$；$b_T = \dfrac{\alpha}{1-\alpha}(\tilde{y}_T - \tilde{\tilde{y}}_T)$；$T$是时间序列的最末期。

可见，二次指数平滑的预测值具有以a_T为截距，b_T为斜率的线性趋势。因此，二次指数平滑适用于有线性趋势的序列。

（3）Holt-Winters 非季节性模型（Holt-Winters-No Seasonal）

该模型与二次指数平滑方法类似，不过有两个平滑系数α和β（$0 \le \alpha, \beta \le 1$）。平滑后的序列$\tilde{y}_t$为：

$$\tilde{y}_{t+k} = a_t + b_t k \quad 对于所有\ k \ge 1 \tag{3.14}$$

其中：

$$a_t = \alpha y_t + (1-\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \tag{3.15}$$

其中：a_t表示截距；b_t表示斜率。可以看出它们都是通过平滑计算得到，需要用简单的方法给出它们的初值。

Holt-Winters 非季节性模型的预测公式为：

$$\tilde{y}_{T+k} = a_T + b_T k \quad 对于所有\ k \ge 1 \tag{3.16}$$

其中：T 是时间序列的最末期。可见，该模型与二次指数平滑方法一样，适用于预测有线性趋势的序列。

（4）Holt-Winters 加法模型（Holt-Winters-Additive）

Holt-Winters 加法模型与 Holt-Winters 非季节性模型相比，主要的不同是加入了加法模型季节因子。该模型有三个平滑系数α、β和γ（$0 \le \alpha, \beta, \gamma \le 1$）。平滑之后的序列$\tilde{y}_t$为：

$$\tilde{y}_{t+k} = a_t + b_t k + S_{t+k} \quad 对于所有\ k \geq 1 \tag{3.17}$$

其中：

$$a_t = \alpha(y_t - S_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1})$$
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$
$$S_t = \gamma(y_t - a_t) + (1-\gamma)S_{t-s} \tag{3.18}$$

其中：a_t 表示截距；b_t 表示斜率；$a_t + b_t k$ 表示趋势；S_t 是加法模型季节因子；s 表示季节周期长度，季度数据 s=4，月度数据 s=12。需要用简单的方法给出季节因子的第一年的初值，以及截距和斜率的初值。

Holt-Winters 加法模型的预测公式为：

$$\tilde{y}_{T+k} = a_T + b_T k + S_{T+k-s} \quad 对于所有\ k \geq 1 \tag{3.19}$$

其中：S_{T+k-s} 使用样本数据最后一年的季节因子。可以看出，Holt-Winters 加法模型适用于既有线性趋势又有加法季节变化的序列。

（5）Holt-Winters 乘法模型（Holt-Winters-Multiplicative）

Holt-Winters 乘法模型是将 Holt-Winters 加法模型中的加法模型季节因子换成乘法模型季节因子。该模型也有三个平滑系数 α 、β 和 γ（$0 \leq \alpha, \beta, \gamma \leq 1$）。平滑之后的序列 $\tilde{y}_t$ 为：

$$\tilde{y}_{t+k} = (a_t + b_t k)S_{t+k} \quad 对于所有\ k \geq 1 \tag{3.20}$$

其中：

$$a_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(a_{t-1} + b_{t-1})$$
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$
$$S_t = \gamma \frac{y_t}{a_t} + (1-\gamma)S_{t-s} \tag{3.21}$$

其中：a_t 表示截距；b_t 表示斜率；$a_t + b_t k$ 表示趋势；S_t 为乘法模型季节因子；s 表示季节周期长度，季度数据 s=4，月度数据 s=12。需要用简单的方法给出季节因子的第一年的初值，以及截距和斜率的初值。

Holt-Winters 乘法模型的预测公式为：

$$\tilde{y}_{T+k} = (a_T + b_T k)S_{T+k-s} \quad 对于所有\ k \geq 1 \tag{3.22}$$

其中：S_{T+k-s} 使用样本数据最后一年的季节因子。可以看出，Holt-Winters 乘法模型适用于既有线性趋势又有乘法季节变化的序列。

以上介绍的 5 种指数平滑方法虽然都有各自适用的数据对象，但是它们的总体思想都是利用加权移动平均的方法对序列进行平滑，因此主要反映的都是近期数据

的变化，适用于短期预测。

【实验数据】

我国流通中现金（M_0）月度数据，样本范围 1996 年 1 月到 2007 年 12 月（参见资料光盘：数据集/平滑技术和季节调整数据/流通中现金（M_0）月度数据.xls）。数据来源为中经网统计数据库。

【实验内容】

M_0是指流通领域中以现金形式存在的货币，它是流动性最强的金融资产，是最活跃的货币，也是中央银行关注和调节的重要目标。它与一国的经济发展水平、消费物价水平以及人们的支付习惯有着密切的关系，因此表现出一定的趋势特征和波动规律。本次实验，同学们可以根据我国流通中现金（M_0）1996 年 1 月到 2007 年 12 月的数据，分别运用简单移动平均、季节调整和指数平滑来寻找发现 M_0 的趋势特征和波动规律，最后运用指数平滑技术对 2008 年 1 月到 3 月 M_0 的值进行预测。

【实验步骤】

1. 根据数据频率和时间范围，创建 Eviews 工作文件（Workfile）。
2. 录入数据，并对序列进行初步分析。分别绘制 M_0 原始序列和 M_0 对数序列的折线图，分析序列的基本趋势。
3. 运用简单移动平均对 M_0 对数序列进行平滑。分别做 3 期和 7 期简单移动平均。同时，利用组形式的折线图进行对数序列与 3 期、7 期简单移动平均序列比较，体会 *n* 期简单移动平均中，*n* 的大小与平滑程度的关系。
4. 对 M_0 对数序列进行季节调整。根据 M_0 对数序列的趋势特征和季节波动规律，选择对应的季节调整模型，得到季节因子。观察季节调整前后的序列图，分析季节

调整的作用。

5. 指数平滑。(1)平滑 M_0 对数序列。根据 M_0 对数序列的趋势特征和波动规律，选择对应的指数平滑模型，对 M_0 对数序列进行指数平滑。并得到对应的指数平滑预测公式。(2)预测 M_0 对数序列。利用上面选择的指数平滑模型，预测 M_0 对数序列和 M_0 原始序列 2008 年 1 月到 3 月的值。并根据预测值和预测效果图评价预测效果。

6. 综合上述实验步骤得出的结果，得出最终结论。总结实验过程中的问题以及得到的经验教训，完成实验报告。

【问题思考】

1. 为什么用 n 期简单移动平均方法进行序列平滑时，n 越大，平滑的程度越高？
2. M_0 的季节特性表明我国居民什么样的支付习惯？
3. 各种指数平滑模型分别适用于具有什么特征的时间序列？
4. M_0 的增长和波动规律是什么样的？应该应用哪种指数平滑模型进行预测？

【实验总结】

通过本实验，加深对平滑技术和季节调整模型形式和基本原理的理解和掌握。熟练运用加法模型和乘法模型对样本序列进行季节调整。掌握运用指数平滑方法进行外推预测的基本方法。写出实验报告。

实验四　随机时间序列的特征观察

【实验目的与要求】

1. 准确掌握随机过程的平稳性和非平稳性的特征和判断方法。
2. 熟练掌握运用相关分析图判断随机过程是否平稳。
3. 学会利用相关分析图判断序列的季节性。
4. 熟练掌握运用相关分析图判断齐次非平稳序列平稳化需要的差分次数。
5. 在教师的指导下独立完成实验，得出正确的结果，并完成实验报告。

【实验准备知识】

1. 随机时间序列简介

随机时间序列，指序列 $y_1, y_2, \cdots, y_T$ 的每个数值都是从一个概率分布中随机得到的，即该时间序列是由某个随机过程生成的。如果我们能够描绘出它的随机特征并构造模型，将有利于我们对序列未来可能值的概率进行推断。

一般来说，我们可假定序列 $y_1, y_2, \cdots, y_T$ 的每个数值都是从一组联合分布的随机变量中获得。如果我们可数量化地确定序列的概率分布，就可以确定序列未来数值的概率。

然而，我们不可能完全确定时间序列的概率分布，但通常情况下可以构造一个简单的时间序列模型，以便解释它的随机性并用于预测目的。模型不必（一般也不

会）与序列的过去实际行为完全一样，因为序列和模型都是随机的，只要模型能够刻画序列的随机特征即可。

2. 平稳和非平稳时间序列

如果随机过程的随机特征不随时间变化，则过程是平稳的；相反，如果随机过程的特征随时间变化，则过程是非平稳的。如果一个过程是非平稳的，用一个简单的代数模型来反映时间序列的过去和未来通常十分困难。如果一个过程是平稳的，则可用具有确定系数的方程来将时间序列模型化，且方程的系数可以利用序列的过去数据估计得到。这有点类似于单方程回归模型，其中解释变量与被解释变量相关，在假定方程所描绘的结构关系不随时间改变的前提下，估计出系数。如果结构关系随时间变化，就不能应用回归的模型技术来进行预测。

（1）平稳过程的性质

任意随机时间序列 $y_1, y_2, \cdots, y_T$ 的每个数值都可以认为是由一组联合分布的随机变量生成，联合概率分布函数假定为 $p(y_1, \cdots, y_T)$。一个未来的观测 y_{T+1} 可以被认为是由条件概率分布函数 $p(y_{T+1} \mid y_1, \cdots, y_T)$ 生成。我们定义平稳过程为其联合分布和条件分布均为不随时间而变化的过程，即如果序列 y_t 是平稳的，则对任意的 t，k 和 m，都有联合概率分布函数：

$$p(y_t, \cdots, y_{t+k}) = p(y_{t+m}, \cdots, y_{t+k+m}) \tag{4.1}$$

且 $$p(y_t) = p(y_{t+m}) \tag{4.2}$$

如果序列 y_t 是平稳的，则序列的数学期望 $\mu_y = E(y_t)$ 是稳定的，即对于任意的 t 和 m，总有 $E(y_t) = E(y_{t+m})$；序列的方差 $\sigma_y^2 = E[(y_t - \mu_y)^2]$ 是稳定的，即对于任意的 t 和 m，总有 $E[(y_t - \mu_y)^2] = E[(y_{t+m} - \mu_y)^2]$；对于任意滞后 k，序列的协方差 $\gamma_k = Cov(y_t, y_{t+k}) = E[(y_t - \mu_y)(y_{t+k} - \mu_y)]$ 是稳定的，即对于任意的 t 和 m，总有 $Cov(y_t, y_{t+k}) = Cov(y_{t+m}, y_{t+m+k})$。此时，随机过程的数学期望 μ_y 可由序列的样本均值 $\overline{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$ 估计；方差 σ_y^2 可由序列的样本方差 $\hat{\sigma}_y^2 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \overline{y})^2$ 估计。

（2）白噪声过程——一个典型的平稳随机过程

白噪声（White Noise），是数理统计和经济计量中经常提到的一个概念，它是指这样的一个随机过程：

$$y_t = u_t \tag{4.3}$$

其中：u_t 来自均值为 0 的独立同分布，即 u_t~IID(0，σ^2)。

可知，对于任意的 t 和 m，随机序列 y_t 的数学期望 $\mu_y = E(y_t) = E(y_{t+m}) = 0$；方差 $E[(y_t - \mu_y)^2] = E[(y_{t+m} - \mu_y)^2] = \sigma^2$；由于是独立同分布，协方差 $Cov(y_t, y_{t+k}) = Cov(y_{t+m}, y_{t+m+k}) = 0$。因此，白噪声过程是一个典型的平稳随机过程。图 4—1 为一个 u_t~IID(0，1)的白噪声过程的序列图。

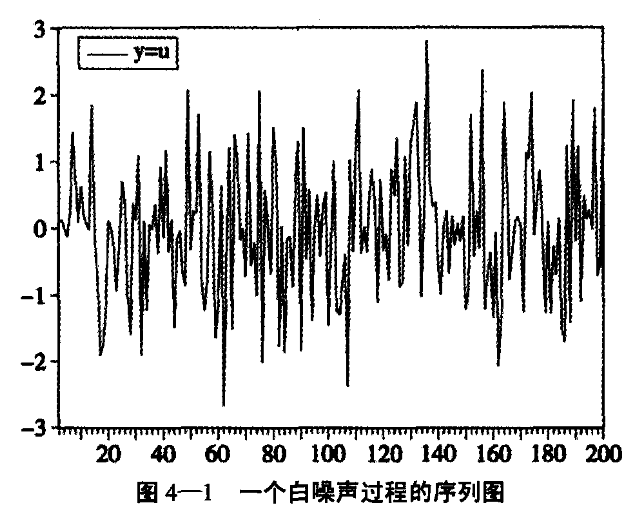


图 4—1　一个白噪声过程的序列图

（3）随机游走过程——一个典型的非平稳随机过程

随机游走（random walk），是由下式定义的一个随机过程：

$$y_t = y_{t-1} + u_t,\ u_t \sim \text{IID}(0,\ \sigma^2) \tag{4.4}$$

可见，y_t 的每一次变化均来自于一个均值为零的独立同分布。易知，y_t 的数学期望 $E(y_t) = E(y_{t-1}) = \cdots = y_0$；又由于 u_t~ IID(0，σ^2)，y_t 的方差 $\sigma_y^2 = E(y_t^2)$ $= E[(y_{t-1} + u_t)^2] = E(y_{t-1}^2) + \sigma^2 = E(y_{t-2}^2) + 2\sigma^2 = \cdots = E(y_{t-n}^2) + n\sigma^2$，从这个迭代可以看出，方差可以无限大，因而无法定义。因此，随机游走过程是一个典型的非平稳随机过程。图 4—2 为一个 $y_0 = 0$，u_t~IID(0，1)的随机游走过程的序列图。

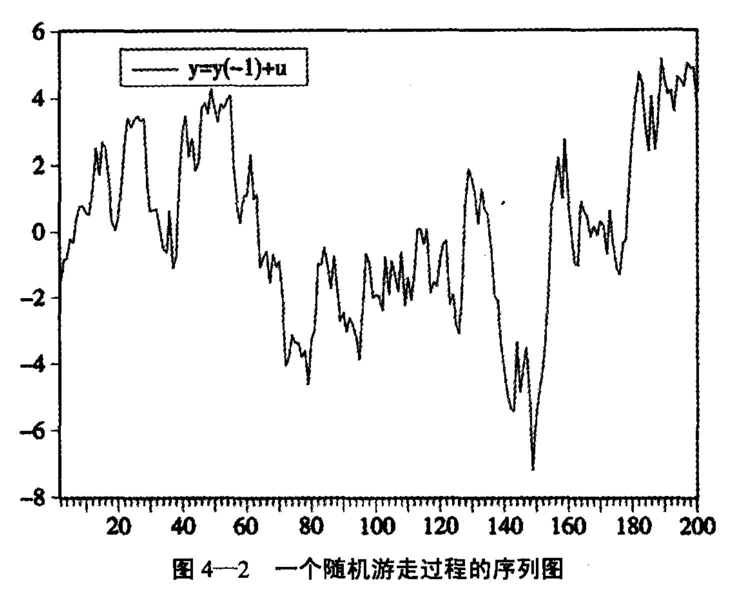


图 4—2　一个随机游走过程的序列图

（4）齐次非平稳过程

我们在实际中遇到的时间序列可能只有极少数属于平稳序列。然而幸运的是，我们遇到的大多数非平稳时间序列（包括在经济、金融和商业中遇到的多数非平稳时间序列）有很好的特性，即可以通过一次或多次差分后成为平稳序列。

我们将这样的非平稳序列称作齐次非平稳过程。原序列变换为平稳序列所需要的差分次数称作齐次的阶数。因此，如果 y_t 是一阶齐次的非平稳过程。则序列

$$w_t = y_t - y_{t-1} = \Delta y_t \tag{4.5}$$

就是平稳的。

如果 y_t 是二阶齐次的非平稳过程。则序列

$$w_t = \Delta^2 y_t = \Delta y_t - \Delta y_{t-1} \tag{4.6}$$

就是平稳的。

很明显，简单随机游走过程 $y_t = y_{t-1} + u_t$，u_t~IID(0，σ^2)就是一阶齐次的非平稳过程。而随机游走过程的一次差分序列

$$w_t = y_t - y_{t-1} = \Delta y_t = u_t \tag{4.7}$$

因为 u_t~IID(0，σ^2)，显然 w_t 是一个平稳过程，事实上 w_t 就是一个白噪声过程。

3. 刻画时间序列的自相关系数和偏自相关系数

（1）自相关系数

自相关系数告诉我们在序列 y_t 的邻近数据点之间存在多大程度的相关性，它可

以部分地刻画随机过程的性质。我们定义滞后期为 k 的自相关系数为：

$$\rho_k=\frac{E[(y_t-\mu_y)(y_{t+k}-\mu_y)]}{\sqrt{E[(y_t-\mu_y)^2]E[(y_{t+k}-\mu_y)^2]}}=\frac{Cov(y_t,y_{t+k})}{\sigma_{y_t}\sigma_{y_{t+k}}} \tag{4.8}$$

对于平稳过程，方差是稳定的，所以可变为：

$$\rho_k=\frac{E[(y_t-\mu_y)(y_{t+k}-\mu_y)]}{\sigma_y^2} \tag{4.9}$$

分子恰是 y_t 和 y_{t+k} 的自协方差 γ_k，所以 $\rho_k=\frac{\gamma_k}{\sigma_y^2}=\frac{\gamma_k}{\gamma_0}$。可见，$\rho_0=1$ 对任何随机过程都成立。白噪声的自相关函数为：$\rho_0=1$ 且对于 $k>0$，$\rho_k=0$。

实际中，我们需要估计自相关系数，即所谓样本自相关系数：

$$r_k=\frac{\sum_{t=1}^{T-k}(y_t-\overline{y})(y_{t+k}-\overline{y})}{\sum_{t=1}^{T}(y_t-\overline{y})^2} \tag{4.10}$$

其中：r_k 是自相关系数 ρ_k 的估计；$\overline{y}$ 是样本均值；T 是样本期数。

与简单相关系数一样，样本自相关系数 r_k 的取值范围是[-1，1]，并且 $|r_k|$ 越接近1，自相关程度越高。

（2）偏自相关系数

偏自相关系数是指对于时间序列 y_t，在给定 $y_{t-1},y_{t-2},\cdots,y_{t-k+1}$ 的条件下，y_t 与 y_{t-k} 之间的条件相关系数，用 $\varphi_{k,k}$ 表示，有 $-1\leqslant\varphi_{k,k}\leqslant1$。

$$\varphi_{k,k}=\begin{cases} r_1 & k=1 \\ \dfrac{r_k-\sum_{j=1}^{k-1}\varphi_{k-1},r_{k-j}}{1-\sum_{j=1}^{k-1}\varphi_{k-1},r_{k-j}} & k=2,3,\cdots \end{cases} \tag{4.11}$$

其中：r_k 是滞后 k 期的自相关系数。

偏自相关系数的一致估计为：

$$\varphi_{k,j}=\varphi_{k-1,j}-\varphi_{k,k}\varphi_{k=1,k-j} \qquad j=1,2,\cdots,k-1 \tag{4.12}$$

（3）Q 统计量检验

Box-Pierce (1970) 提出的 Q 统计量，可以用来检验序列的相关性，还可以用来

检验残差序列是否为白噪声。Q 检验的零假设是：

$$H_0:\quad \rho_1 = \rho_2 = \cdots = \rho_K = 0$$

即序列不存在自相关，是一个白噪声过程。其中，ρ_i 表示自相关系数。Q 统计量定义为：

$$Q = T\sum_{k=1}^{K} r_k^{\,2} \tag{4.13}$$

其中：T 表示样本容量；r_k 表示样本自相关系数值；K 表示自相关系数的个数，即最大滞后期。若样本较大，K 可取[T/10]或者[$\sqrt{T}$]；若样本较小，K 可取[T/4]。当 $T\to\infty$ 时，Q 渐近服从 $\chi^2(K)$ 分布。

Ljung 和 Box 认为（4.14）式定义的 Q 统计量的分布与 $\chi^2(K)$ 分布存在差异，于是提出修正的 Q 统计量。

$$Q = T(T+2)\sum_{k=1}^{K}\frac{r_k^{\,2}}{T-k} \tag{4.14}$$

其中：r_k，K，T 的定义如（4.14）式。

修正的 Q 统计量渐近服从 $\chi^2(K)$ 分布。且它的近似性比原 Q 统计量的近似性更好。显然若序列不是白噪声，序列的自相关系数不等于零，则 Q 值将很大，反之 Q 值将很小。

判别规则是：若 $Q<\chi^2_\alpha(K)$，则接受 H_0；

若 $Q>\chi^2_\alpha(K)$，则拒绝 H_0。

其中：α 表示检验水平。

这里需要注意，自相关性和平稳性是两个概念，是否存在自相关性并不能说明序列是否平稳。

（4）刻画时间序列的相关分析图

通常情况下，我们将序列的自相关系数和偏自相关系数绘制成图，并标出一定的随机区间，称为相关分析图。比如某个白噪声过程 $y_t = u_t$，u_t~IID(0，1)的相关分析图如图 4—3 所示。该图左半部分为序列的自相关和偏自相关系数图，虚线为 5% 显著性水平下的置信带，右半部分的五列数值分别为：滞后期 K、自相关系数 r_k、

偏自相关系数$\varphi_{k,k}$、Q 统计量和相伴概率。可以看出，该白噪声过程的自相关系数几乎都落入虚线内，表示与 0 没有显著差异。Q 统计量的相伴概率都大于 0.05，可以接受没有自相关的原假设。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.009	0.009	0.0170	0.896
		2	0.017	0.017	0.0733	0.964
		3	0.102	0.102	2.2077	0.530
		4	-0.024	-0.026	2.3233	0.677
		5	-0.023	-0.026	2.4305	0.787
		6	-0.017	-0.026	2.4872	0.870
		7	-0.016	-0.009	2.5384	0.924
		8	-0.082	-0.078	3.9626	0.860
		9	0.034	0.040	4.2046	0.897
		10	0.011	0.014	4.2285	0.936
		11	-0.095	-0.083	6.1337	0.864
		12	0.033	0.022	6.3603	0.897
		13	0.007	0.005	6.3698	0.932
		14	-0.102	-0.090	8.6238	0.854
		15	0.093	0.088	10.488	0.788

图 4—3　某个白噪声过程的相关分析图

再比如某个随机游走过程 $y_t = y_{t-1} + u_t$，$y_0 = 0$，u_t~IID(0，1)的相关分析图如图 4—4 所示。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.924	0.924	172.36	0.000
		2	0.845	-0.055	317.41	0.000
		3	0.761	-0.082	435.56	0.000
		4	0.683	-0.006	531.11	0.000
		5	0.623	0.081	611.04	0.000
		6	0.572	0.019	678.82	0.000
		7	0.526	-0.011	736.46	0.000
		8	0.483	-0.008	785.36	0.000
		9	0.433	-0.067	824.85	0.000
		10	0.375	-0.081	854.54	0.000
		11	0.332	0.085	877.97	0.000
		12	0.301	0.056	897.33	0.000
		13	0.285	0.056	914.77	0.000
		14	0.263	-0.081	929.71	0.000
		15	0.236	-0.045	941.82	0.000

图 4—4　某个随机游走过程的相关分析图

由图 4—4 可以看出，该随机游走过程的自相关系数随着 K 增加缓慢下降，但一直到 15 期也没有落入置信区间。Q 统计量的相伴概率都小于 0.05，可以确定序列存在自相关。

（5）平稳性和自相关系数

我们可通过考察自相关系数图来初步确定一个序列是否平稳或确定一个齐次非平稳序列需要差分多少次才能平稳。平稳序列的自相关系数都为 0（如图 4—3 所示），或者随着滞后期 K 的增加而迅速下降为 0（如图 4—5 所示）。非平稳序列一般却不是这样（如图 4—4 所示），其自相关系数是缓慢衰减的，并长期与 0 有显著差异。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.064	-0.064	0.4512	0.502
		2	-0.339	-0.345	13.237	0.001
		3	0.029	-0.026	13.328	0.004
		4	0.173	0.065	16.732	0.002
		5	-0.129	-0.121	18.628	0.002
		6	-0.031	0.031	18.742	0.005
		7	0.069	-0.008	19.290	0.007
		8	-0.028	-0.045	19.385	0.013
		9	0.105	0.170	20.685	0.014
		10	-0.002	-0.023	20.685	0.023
		11	0.011	0.103	20.701	0.037
		12	0.069	0.114	21.280	0.046
		13	0.068	0.081	21.849	0.058
		14	-0.051	0.068	22.174	0.075
		15	-0.136	-0.120	24.502	0.057

图 4—5　某个平稳过程的相关分析图

4. 季节性和自相关系数

如前所述，季节性是以年为基准的周期性的行为。季节性的高峰和低潮通常可以通过直接观察时间序列而得到。然而，如果时间序列波动得厉害，季节性的高峰和低潮就不一定能从其他波动中区分开来。季节性的识别是很重要的，因为它提供了时间序列规律性的信息。这个规律可帮助我们对序列进行预测。幸运的是借助自相关系数可以比较容易地对季节性进行识别。

如果月度时间序列 y_t 有年度的季节周期性，则序列的数据将显示每 12 期与它的前 12 期或滞后 12 期的一定程度的相关性。换言之，我们应该能够看到 y_t 和 y_{t-12} 有某种程度的相关性。因为 y_t 和 y_{t-12} 相关，y_{t-12} 和 y_{t-24} 也相关，所以 y_t 和 y_{t-24} 相关，类似的 y_t 和 y_{t-36} 等都相关。于是，样本自相关系数 r_k，将在 k=12，24，36，48 等等明显地反映出这些相关性。因此，我们可以通过观察自相关系数的有规律的峰值来识别季节性，甚至可识别那些由时间序列本身无法辨别的季节性峰值。当我们考察月度数据，季节周期一般为 12 个月，因此自相关系数的最大滞后阶数 k 取 12，24，36 等，如果是季度数据，一般就取 4，8，12 等。

以实验三中我国流通中现金（M_0）月度数据为例，图 4—6 为 M_0 序列折线图，可以看出每年一月份有明显的峰值。

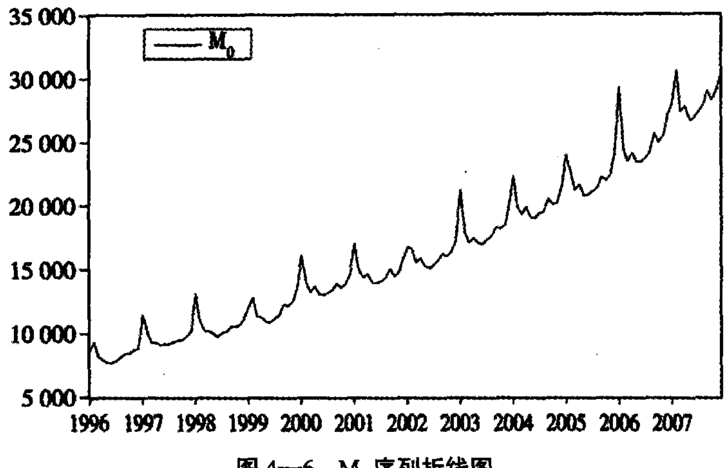


图 4—6　M_0 序列折线图

M_0 序列的相关分析图如图 4—7 所示，可以看出 M_0 序列明显非平稳，但反映不出季节性，这是由于序列的强趋势性掩盖了季节性。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.957	0.957	134.69	0.000
		2	0.926	0.112	261.52	0.000
		3	0.900	0.076	382.35	0.000
		4	0.872	-0.026	496.54	0.000
		5	0.846	0.016	604.87	0.000
		6	0.823	0.014	707.96	0.000
		7	0.804	0.065	807.22	0.000
		8	0.788	0.040	903.26	0.000
		9	0.768	-0.034	995.19	0.000
		10	0.750	0.002	1083.4	0.000
		11	0.730	-0.023	1167.7	0.000
		12	0.716	0.058	1249.3	0.000
		13	0.687	-0.158	1325.2	0.000
		14	0.657	-0.072	1394.9	0.000
		15	0.635	0.044	1460.6	0.000

图 4—7　M_0 序列的相关分析图

为了剔除趋势性，我们对 M_0 序列做一次差分，然后对差分序列绘制自相关系数图，如图 4—8 所示。可以看出差分后的 M_0 序列虽然平稳，但样本自相关系数 r_k，在 k=12，24，36 有明显的峰值，这充分反映出季节性。

Autocorrelation Partial Correlation

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36

图 4—8 M_0序列差分后的自相关系数图

【实验数据】

我国金融机构 2000 年 1 月到 2007 年 12 月财政存款余额数据（参见资料光盘：数据集/随机时间序列的特征观察数据/我国金融机构财政存款余额月度数据.xls）。数据来源为中经网统计数据库。

【实验内容】

财政存款，又称财政金库存款，指财政部门存在银行的货币资金。中国由中国人民银行代理国家金库，国家授权中国人民银行办理财政收入和支出的出纳业务。财政先收后支的待用款和收大于支的结余款在银行形成财政存款。财政存款是中国人民银行信贷资金的一项重要来源，各专业银行吸收的财政存款都必须上缴给中国人民银行使用。本次实验，同学们可以根据我国金融机构 2000 年 1 月到 2007 年 12 月财政存款余额数据，利用相关分析图观察随机序列的特征和季节性，并且判断序列是否为齐次非平稳序列。如果是，判断平稳化需要的差分次数。

【实验步骤】

1. 根据数据频率和时间范围，创建 Eviews 工作文件（Workfile）。
2. 录入数据，并对序列进行初步分析。绘制金融机构财政存款余额原始序列和对数序列的折线图，分析序列的基本趋势。
3. 利用相关分析图观察对数序列的随机特征。分别观察对数序列和一次差分后的对数序列的相关分析图，判断序列平稳性和季节性。
4. 对财政存款余额对数序列进行季节调整，再观察相关分析图。分别观察季节调整后的对数序列和一次差分后序列的相关分析图，判断序列平稳性和平稳化需要的差分次数。
5. 综合上述实验步骤得出的结果，得出最终结论。总结实验过程中的问题以及得到的经验教训，完成实验报告。

【问题思考】

1. 平稳序列和非平稳序列的相关分析图各有什么特征？

2. 如何利用相关分析图判断序列的季节性？
3. 齐次非平稳序列平稳化需要的差分次数如何判断？

【实验总结】

通过本实验，加深对平稳序列和非平稳序列特征的理解和掌握。熟练运用相关分析图判断样本序列的平稳性。学会利用相关分析图判断样本序列的季节性。掌握齐次非平稳序列平稳化需要差分次数的判断方法。写出实验报告。

实验五　单位根检验、协整与误差修正模型

【实验目的与要求】

1. 准确掌握单位根检验方程的形式和检验原理。
2. 准确掌握单整、协整和误差修正模型的概念和形式。
3. 学会利用单位根检验方法对样本序列进行平稳性检验和协整关系检验。
4. 熟练掌握运用误差修正模型对样本序列之间的短期、长期关系进行分析。
5. 在教师的指导下独立完成实验，得出正确的结果，并完成实验报告。

【实验准备知识】

在上个实验中，我们学习了如何运用相关分析图判断随机过程是否平稳，但这种方法比较粗略。检验随机过程是否平稳的一种比较正式的方法就是单位根检验。

在介绍单位根检验之前，我们有必要认识几种典型的非平稳随机过程。

1. 几种典型的非平稳随机过程

（1）随机游走过程

$$y_t = y_{t-1} + u_t,\ u_t \sim \text{IID}(0,\ \sigma^2) \tag{5.1}$$

随机游走过程上个实验已经介绍，这里不再赘述。图 5—1 为一个 $y_0=0$，$u_t\sim$ IID（0，1）的随机游走过程的序列图。

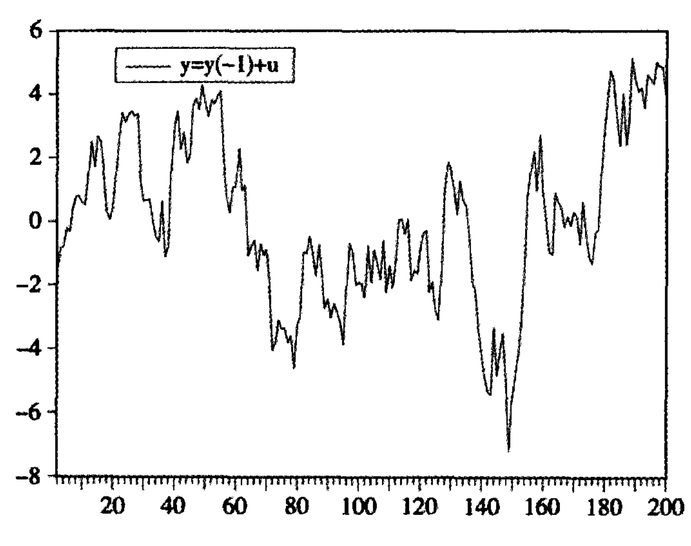


图 5—1 一个随机游走过程的序列图

（2）随机趋势过程

$$y_t=y_{t-1}+a+u_t\text{，}\quad u_t\sim\text{IID（0，}\sigma^2\text{）}\tag{5.2}$$

其中：a 称作位移项或漂移项。将上式作如下迭代变换：

$$y_t=y_{t-1}+a+u_t=(y_{t-2}+a+u_{t-1})+a+u_t=\cdots=at+y_0+\sum_{i=1}^{t}u_i\tag{5.3}$$

可知，y_t 由时间趋势项 at 和 $y_0+\sum_{i=1}^{t}u_i$（可看做截距项）组成。在不存在任何冲击 u_t 的情况下，截距项为 y_0，而每个冲击 u_t 都表现为截距的移动。每个冲击 u_t 对截距项的影响都是持久的，导致序列的条件均值发生变化，所以称这样的过程为随机趋势过程或有漂移项的随机游走过程。图 5—2 为一个 $y_t=y_{t-1}+0.3+u_t$，$y_0=0$，$u_t\sim$ IID（0，1）的随机趋势过程的序列图。

图 5—2 表明，虽然总趋势不变，但该过程围绕趋势项上下游动。由（5.3）式还可以看出，a 是时间趋势项的系数（原序列的增长速度）。a 为正时，趋势向上；a 为负时，趋势向下。

（3）趋势非平稳过程

$$y_t=y_{t-1}+a+\delta t+u_t\text{，}\quad u_t\sim\text{IID（0，}\sigma^2\text{）}\tag{5.4}$$

其中：a 称作位移项或漂移项；δt 称作趋势项。可见，趋势非平稳过程是随机趋势和确定性趋势的混合随机过程。将上式作如下迭代变换：

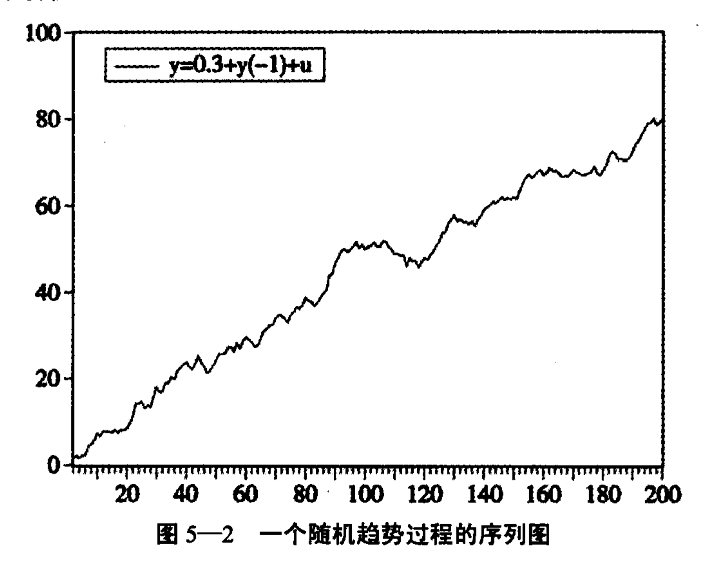


图 5—2 一个随机趋势过程的序列图

$$
\begin{aligned}
y_t &= y_{t-1}+a+\delta t+u_t=(y_{t-2}+a+\delta(t-1)+u_{t-1})+a+\delta t+u_t \\
&=\cdots=y_0+at+\delta[t+(t-1)+\cdots+2+1]+\sum_{i=1}^{t}u_i \qquad (5.5)\\
&=y_0+at+\frac{\delta}{2}t(t+1)+\sum_{i=1}^{t}u_i=(a+\frac{\delta}{2})t+\frac{\delta}{2}t^2+y_0+\sum_{i=1}^{t}u_i
\end{aligned}
$$

由（5.5）式可以看出，趋势非平稳过程的趋势项中包括 t 的 1 次和 2 次项。图 5—3 为一个 $y_t=y_{t-1}+0.3+0.1t+u_t$， $y_0=0$， $u_t\sim$ IID（0，1）的趋势非平稳过程的序列图。

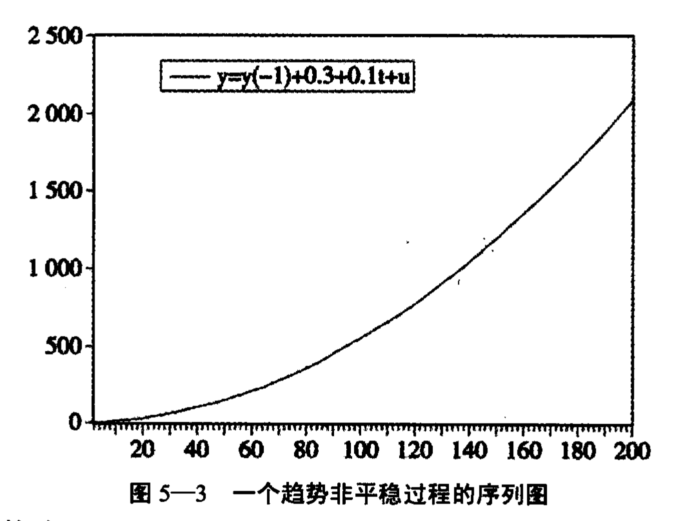


图 5—3 一个趋势非平稳过程的序列图

2. 单位根检验

（1）DF 检验

考虑三个随机过程：

$$y_t = \rho y_{t-1} + u_t \tag{5.6}$$

$$y_t = \rho y_{t-1} + a + u_t \tag{5.7}$$

$$y_t = \rho y_{t-1} + a + \delta t + u_t \tag{5.8}$$

其中：ρ 是有理数；a 是常数项；δt 是时间趋势项；u_t 是白噪声。若 $|\rho|<1$，则序列 y_t 是平稳的；若 $|\rho|=1$，则序列 y_t 是非平稳的（即（5.1）、（5.2）、（5.4）式）；而当 $|\rho|>1$，序列 y_t 是强非平稳的，是爆炸性的，没有实际意义。因此，检验 y_t 平稳性，我们要检验的就是 $|\rho|$ 是否严格小于 1。

实际检验时，我们将（5.6）、（5.7）、（5.8）式左右同时减去 y_{t-1} 得到三个检验方程为：

$$\Delta y_t = \gamma y_{t-1} + u_t \tag{5.9}$$

$$\Delta y_t = \gamma y_{t-1} + a + u_t \tag{5.10}$$

$$\Delta y_t = \gamma y_{t-1} + a + \delta t + u_t \tag{5.11}$$

其中：$\gamma = \rho - 1$。检验假设为：

$H_0 : \gamma = 0$　（y_t 非平稳）

$H_1 : \gamma < 0$　（y_t 平稳）

参数 γ 估计值的显著性检验的 t 统计量不服从常规的 t 分布，Dickey 和 Fuller 于 1979 年给出了检验用的模拟临界值，故该检验称为 DF 检验。Eviews 中使用的是 Mackinnon 改进的单位根检验临界值。DF 检验做的是左单端检验，检验规则是：若 DF>临界值，则接受 H_0，y_t 非平稳；若 DF<临界值，则拒绝 H_0，y_t 平稳。

（2）ADF 检验

DF 检验适用于序列为 AR(1)过程。如果序列存在高阶滞后相关，就会破坏随机扰动项是白噪声的假设，这时可使用扩展的 DF 检验（ADF 检验）来检验含有高阶序列相关的序列的单位根。除了检验方程不同外，ADF 检验的检验假设、检验规则等都与 DF 检验类似。

ADF 检验是假设序列 y_t 为 AR(p)过程。检验方程为：

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^{p} \Delta y_{t-i} + u_t \tag{5.12}$$

$$\Delta y_t = \gamma y_{t-1} + a + \sum_{i=1}^{p} \Delta y_{t-i} + u_t \tag{5.13}$$

$$\Delta y_t = \gamma y_{t-1} + a + \delta t + \sum_{i=1}^{p} \Delta y_{t-i} + u_t \tag{5.14}$$

ADF 检验中很重要的问题是滞后阶数 p 的选择，通常采用 AIC 准则（Akaike Information Criterion）来确定。

不论是 DF 检验还是 ADF 检验都有三种检验方程，选择哪种形式也很重要，即是否加入常数项和时间趋势项。我们可以通过的观察序列的折线图来判断，近似于随机游走的序列（类似图 5—1 形式）可采用（5.9）、（5.12）式进行检验，即没有添加项；有线性趋势的序列（类似图 5—2 形式）可采用（5.10）、（5.13）式进行检验，即加入常数项；有二次趋势的序列（类似图 5—3 形式）可采用（5.11）、（5.14）式进行检验，即同时加入常数项和时间趋势项。当通过直观观察折线图也很难判断检验方程形式的时候，可以利用辅助方程，观察常数项和时间趋势项系数的显著性，从而判断检验方程的形式。

3. 单整、协整与误差修正模型

（1）单整

我们在实验四中介绍的齐次非平稳过程虽然是非平稳的，但是可以通过一次或多次差分后成为平稳序列。像这种非平稳序列可以通过差分运算而平稳，我们称这样的序列为单整（integration）序列。

严格的定义为：如果序列 y_t 通过 d 次差分达到平稳，而这个序列 $d-1$ 次差分不平稳，则称序列 y_t 为 d 阶单整序列，记作 $y_t \sim \boldsymbol{I}$（d）。其中，d 表示单整阶数，是序列包含的单位根个数（使序列平稳而差分的次数）。如果序列 y_t 本身是平稳的，则为零阶单整序列，记作 $y_t \sim \boldsymbol{I}$（0）。

（2）协整

在时间序列分析中，如果用两个独立的非平稳时间序列建立回归模型，往往会得到具有统计显著性的回归参数，这种现象称为虚假回归，在统计上也称为无意义相关。由于实际中的大多数经济时间序列是非平稳的，为了避免虚假回归问题，通常采用差分的方法使序列平稳之后再建立模型，但是变换后的序列限制了所讨论经济问题的范围，有时也不具有直接的经济意义。

1987 年 Engle 和 Granger 提出了协整理论，为非平稳序列建模提供了另一种途径。有些时间序列，虽然本身是非平稳的，但其某种线性组合却是平稳的，这个线性组合反映了变量之间长期稳定的均衡关系，称为协整（cointegration）。协整的严格定义为：

如果 k 维向量 $\mathbf{y}_t$=（$y_{1t}, y_{2t}, \cdots, y_{kt}$）的分量都是 d 阶单整序列，即 I（d）。存在一

个非零向量β=（$\beta_1,\beta_2,\cdots,\beta_k$），使得β$y'_t$~$I$（$d-b$），0<$b$≤$d$，则向量 y_t 的分量间被称为(d，b)阶协整，记为 y_t~CI（d，b），β称作协整向量。

（3）协整检验

这里我们只讨论两个变量的协整关系检验。为检验两个变量 x_t 和 y_t 是否协整，Engle 和 Granger 于 1987 年提出了两步检验法，称为 EG 检验。若序列 x_t 和 y_t 都是 d 阶单整的，用一个变量对另一个变量回归，即：

$$y_t = \alpha + \beta x_t + u_t \tag{5.15}$$

用 $\hat{\alpha}$ 和 $\hat{\beta}$ 表示回归系数的估计值，则模型的残差为：

$$\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta} x_t \tag{5.16}$$

若 $\hat{u}_t$~I（0），则 x_t 和 y_t 有协整关系，即 x_t 和 y_t 有长期稳定的均衡关系，且向量（1，$-\hat{\beta}$）为协整向量，（5.15）式为协整方程。这里，序列 x_t 和 y_t 的 d 阶单整检验和 $\hat{u}_t$ 序列的零阶单整检验，都可以使用前面介绍的单位根检验方法。

（4）误差修正模型（ECM）

首先介绍自回归分布滞后模型（Auto-Regressive Distributed Lag，ADL）。如果一个内生变量 y_t 只被表示成同一时点的外生变量 x_t 的函数，x_t 对 y_t 的长期影响很容易求出。然而如果每个变量的滞后也出现在模型之中，其长期影响将通过分布滞后函数反映，这就是自回归分布滞后模型（ADL）。

考虑一阶自回归分布滞后模型，记为 ADL（1，1）：

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 x_t + \beta_3 x_{t-1} + u_t \tag{5.17}$$

其中：$u_t \sim$ IID（0，σ^2）。移项后，整理可得：

$$\Delta y_t = \beta_0 + (\beta_1 - 1)(y_t - \frac{\beta_2 + \beta_3}{1-\beta_1} x_t)_{t-1} + \beta_2 \Delta x_t + u_t \tag{5.18}$$

方程（5.17）和（5.18）是等价的，但每个方程有不同的解释和意义，方程（5.18）被称为误差修正模型（Error Correction Model，简记为 ECM），其中，$y_t - \frac{\beta_2 + \beta_3}{1-\beta_1} x_t$ 称为误差修正项，记为 ecm。模型（5.18）可以简记为：

$$\Delta y_t = \beta_0 + \alpha ecm_{t-1} + \beta_2 \Delta x_t + u_t \tag{5.19}$$

其中：$\alpha=(\beta_1-1)$。

模型（5.18）解释了被解释变量 y_t 的短期波动 Δy_t 是如何被决定的。一方面，它受到解释变量短期波动 Δx_t 的影响；另一方面，取决于 *ecm*。如果 y_t 和 x_t 有长期稳定的均衡关系，即有 $y^*=bx^*$，*ecm* 可改写为：

$$y^*=\frac{\beta_2+\beta_3}{1-\beta_1}x^* \tag{5.20}$$

可见，ecm 反映了 y_t 关于 x_t 在第 t 期的短期偏离，即在短期波动中偏离它们长期均衡关系的程度。一般来说，由于式（5.18）中 $|\beta_1|<1$，所以（5.19）中误差修正项的系数 $\alpha=(\beta_1-1)<0$，通常称为调整系数或修正系数，表示在第 t-1 期 y_{t-1} 关于 $\frac{\beta_2+\beta_3}{1-\beta_1}x_{t-1}$ 之间偏差的调整速度。可见，当 $y_{t-1}>\frac{\beta_2+\beta_3}{1-\beta_1}x_{t-1}$，$ecm_{t-1}$ 为正，αecm_{t-1} 为负，使 Δy_t 减少，反之亦然，这体现了均衡误差对 y_t 的控制。这样，y_t 就在不断“修正”前期“误差”的过程中变化，使 y_t 与 x_t 的关系始终围绕长期均衡关系 $y^*=\frac{\beta_2+\beta_3}{1-\beta_1}x^*$。

最常用的 ECM 模型的估计方法是 Engle 和 Granger 于 1981 年提出的两步法，其基本思想是：

第一步是求模型：

$$y_t=bx_t+u_t \tag{5.21}$$

的 OLS 估计，又称协整回归，得到 $\hat{b}$ 及残差序列：

$$\hat{u}_t=y_t-\hat{b}x_t \tag{5.22}$$

第二步是用 $\hat{u}_t$ 替换（5.19）式中的 *ecm*，即：

$$\Delta y_t=\beta_0+\alpha\hat{u}_{t-1}+\beta_2\Delta x_t+u_t \tag{5.23}$$

再对（5.23）式进行 OLS 估计，得到参数的估计值，从而得到 ECM 模型。

【实验数据】

我国上证综指和深证综指1998年1月9日到2008年3月7日周收盘价数据(参见资料光盘：数据集/单位根检验、协整与误差修正模型数据/上证综指、深证综指周数据)。数据来源为大智慧软件（下载并整理)。

【实验内容】

上证综指和深证综指是反映我国两大股票交易市场上海证券交易所和深圳证券交易所股票行情的重要指标，两者是否存在长期稳定的相关关系，可以在一定程度上反映出两个市场的相关程度。本次实验，同学们可以根据我国上证综指和深证综指1998年1月9日到2008年3月7日周收盘价数据，利用单位根检验、协整关系检验方法来判断上证综指和深证综指序列是否存在协整关系，即长期稳定的相关关系。如果存在协整关系，还可以运用误差修正模型对两者的短期、长期关系进行分析。

【实验步骤】

1. 根据数据频率和时间范围，创建Eviews工作文件（Workfile)。
2. 录入数据，并对序列进行初步分析。分别绘制上证综指和深证综指序列的折线图以及组形式的折线图，分析序列的基本趋势，以及两者的关系。
3. 运用ADF检验对上证综指和深证综指序列进行单位根检验。分别做原序列和差分序列的单位根检验，并判断单整的阶数。根据序列的折线图选择检验方程形式；根据AIC准则选择ADF检验时的最大滞后阶数p。
4. 对上证综指和深证综指序列进行协整检验。如果上证综指和深证综指序列是同阶单整序列，就对上证综指和深证综指序列进行协整关系检验。

5. 建立误差修正模型。如果上证综指和深证综指序列存在协整关系，则建立误差修正模型，分析两者的短期、长期关系。

6. 综合上述实验步骤得出的结果，得出最终结论。总结实验过程中的问题以及得到的经验教训，完成实验报告。

【问题思考】

1. 什么是虚假回归？什么样的序列间容易出现虚假回归问题？
2. ADF 检验时的最大滞后阶数 p 如何确定？
3. 协整检验的两步检验法是指哪两步？
4. 如何运用误差修正模型对序列间的短期、长期关系进行分析？

【实验总结】

通过本实验，加深对单位根检验方程的形式和检验原理的理解和掌握。熟练运用单位根检验方法对样本序列进行平稳性检验和协整关系检验。掌握运用误差修正模型对样本序列之间的短期、长期关系进行分析的基本方法。写出实验报告。

实验六　ARMA 模型应用

【实验目的与要求】

1. 准确掌握 ARIMA（p，d，q）模型的各种形式和基本原理。
2. 熟练掌握识别 ARIMA（p，d，q）模型中的阶数 p，d，q 的方法。
3. 学会建立及检验 ARIMA（p，d，q）模型的基本方法。
4. 熟练掌握运用 ARIMA（p，d，q）模型对样本序列进行拟合和预测。
5. 在教师的指导下独立完成实验，得出正确的结果，并完成实验报告。

【实验准备知识】

1. 线性随机时间序列模型

线性随机时间序列模型（ARMA 模型），是一类常用的随机时间序列模型，由博克斯（Box）、詹金斯（Jenkins）创立，亦称 B-J 方法。ARMA 模型有三种基本类型：自回归（AR：Auto-Regressive）模型、移动平均（MA：Moving Average）模型和自回归移动平均（ARMA：Auto-Regressive Moving Average）模型。此外，这种线性模型既适用于平稳过程，也适用于可平稳化的非平稳过程（可通过差分一次或多次转化为平稳过程），即单整自回归移动平均（ARIMA：Integrated Auto-Regressive Moving Average）模型。线性随机时间序列模型可表示成含有固定待估参数的方程，反映出模型不随时间变化的随机结构。

（1）自回归模型（AR（p）模型）

p 阶自回归模型，记为 AR（p），模型形式为：

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \delta + u_t \tag{6.1}$$

其中：ϕ_i（i=1,…，p）是自回归参数；δ 是常数；u_t 是白噪声过程。可见，AR（p）模型中 y_t 是由它 p 个滞后期的加权平均，再加当期的随机扰动项构成。

引入滞后算子 L，来表示时间的滞后，比如，$Lx_t = x_{t-1}$，$L^n x_t = x_{t-n}$。AR（p）模型可以表达为：

$$y_t = \phi_1 L y_t + \phi_2 L^2 y_t + \cdots + \phi_p L^p y_t + \delta + u_t \tag{6.2}$$

进一步有：

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) y_t = \delta + u_t \tag{6.3}$$

令 $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$，则模型可写成：

$$\Phi(L) y_t = \delta + u_t \tag{6.4}$$

我们将 $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p = 0$ 称为 AR（p）模型的特征方程。可以证明，AR（p）模型平稳的充要条件是特征方程的根都在单位圆之外。

（2）移动平均模型（MA（q）模型）

q 阶移动平均模型，记为 MA（q），模型形式为：

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} \tag{6.5}$$

其中：μ 为常数；θ_i（i=1，…，q）是移动平均参数；u_t 是白噪声。可见，MA（q）模型中 y_t 是当期随机干扰项及其直到 q 期的滞后随机干扰项的加权平均。引入滞后算子 L，MA（q）模型可以表达为：

$$y_t - \mu = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_p L^p) u_t \tag{6.6}$$

令 $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_p L^p$，则模型可写成：

$$y_t - \mu = \Theta(L) u_t \tag{6.7}$$

由于 u_t 是白噪声，可知 MA（q）模型一定是平稳的。而与移动平均模型相联系的一个重要概念是可逆性。我们将 $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_p L^p = 0$ 称为 MA（q）模型的特征方程。如果特征方程的根都在单位圆之外，则称 $\Theta(L)$ 可逆。运用 $\Theta(L)$ 的

逆运算，式（6.7）可以改写为一个无限阶自回归（AR（∞））的形式：

$$(1+\eta_1 L+\eta_2 L^2+\eta_3 L^3+\cdots)(y_t-\mu)=\mu_t \tag{6.8}$$

其中：$(1+\eta_1 L+\eta_2 L^2+\eta_3 L^3+\cdots)=(1+\theta_1 L+\theta_2 L^2+\cdots+\theta_p L^p)^{-1}=\Theta^{-1}(L)$。

类似的，（6.4）式满足平稳性条件时，可改写为一个无限阶移动平均（MA（∞））的形式：

$$y_t-\delta=(1-\varphi_1 L-\varphi_2 L^2-\varphi_3 L^3-\cdots)u_t \tag{6.9}$$

其中：$(1-\varphi_1 L-\varphi_2 L^2-\varphi_3 L^3-\cdots)=(1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p)^{-1}=\Phi^{-1}(L)$。

（3）混合自回归移动平均模型（ARMA（p，q）模型）

（p，q）阶混合自回归移动平均模型，记为 ARMA（p，q），模型形式为：

$$y_t=\phi_1 y_{t-1}+\phi_2 y_{t-2}+\cdots+\phi_p y_{t-p}+\delta+u_t-\theta_1 u_{t-1}-\theta_2 u_{t-2}-\cdots-\theta_q u_{t-q} \tag{6.10}$$

其中：ϕ_i（i=1，…，p）是自回归参数；θ_i（i=1，…，q）是移动平均参数；δ 是常数；u_t 是白噪声过程。可见，ARMA（p，q）模型是由自回归和移动平均两部分共同构成的随机过程，其中 p，q 分别表示自回归和移动平均部分的最大阶数。引入滞后算子 L，ARMA（p，q）模型可以表达为：

$$\Phi(L)y_t=\delta+\Theta(L)u_t \tag{6.11}$$

ARMA（p，q）模型平稳性条件是方程 $\Phi(L)=0$ 的根都在单位圆之外，可逆性条件是方程 $\Theta(L)=0$ 的根都在单位圆之外。

（4）单整自回归移动平均模型（ARIMA（p，d，q）模型）

首先，引入差分算子Δ。比如，一阶差分 $\Delta x_t=x_t-x_{t-1}=x_t-Lx_t=(1-L)x_t$；高阶差分 $\Delta_k x_t=x_t-x_{t-k}=x_t-L^k x_t=(1-L^k)x_t$；二次差分 $\Delta^2 x_t=\Delta x_t-\Delta x_{t-1}=(1-L)^2 x_t=(1-2L+L^2)x_t=x_t-2x_{t-1}+x_{t-2}$。高阶差分常用于季节性数据的差分，如季度数据的 4 阶差分、月度数据的 12 阶差分等。

我们前面介绍过单整序列可以通过一次或多次差分后成为平稳序列。设序列 y_t 为 d 阶单整序列，即 y_t~I（d），则：

$$w_t=\Delta^d y_t=(1-L)^d y_t \tag{6.12}$$

w_t 是平稳序列，于是我们可以对 w_t 建立 ARMA（p，q）模型。如果 $w_t=\Delta^d y_t$ 且

w_t 是一个 ARMA（p，q）过程，则我们称 y_t 是（p，d，q）单整自回归移动平均模型，简记为 ARIMA（p，d，q）。模型形式为：

$$w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + \cdots + \phi_p w_{t-p} + \delta + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \cdots - \theta_q u_{t-q} \tag{6.13}$$

用差分算子和滞后算子表示，则（6.12）式可写成：

$$\Phi(L)\Delta^d y_t = \delta + \Theta(L)u_t \tag{6.14}$$

其中：$\Phi(L)$ 是一个平稳的自回归算子，即 $\Phi(L)=0$ 的根都在单位圆以外；$\Theta(L)$ 是一个可逆的移动平均算子，即 $\Theta(L)=0$ 的根都在单位圆以外。$\Phi(L)\Delta^d$ 是广义自回归算子。

可见，ARIMA（p，d，q）模型是线性随机时间序列模型的普遍形式，AR（p）模型、MA（q）模型和 ARMA（p，q）模型是它的特殊情况。

2. 线性随机时间序列模型的识别

线性随机时间序列模型的识别，首先是用单位根检验的方法，识别 ARIMA（p，d，q）模型中的单整阶数 d，方法已经在前面做过介绍。然后识别自回归阶数 p，移动平均阶数 q，方法是使用前面介绍的自相关函数和偏自相关函数，下面做一简要介绍。

（1）MA（q）模型的自相关函数和偏自相关函数

MA（q）模型

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} \tag{6.15}$$

的自协方差 γ_k 为：

$$\gamma_k = E(y_{t+k}y_t) = \begin{cases} \sigma^2(1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2), & k=0 \\ \sigma^2(\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{k-q}\theta_q), & 0<k\le q \\ 0, & k>q \end{cases} \tag{6.16}$$

进而得到自相关函数：

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} 1, & k=0 \\ (\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{k-q}\theta_q)/(1+\theta_1^2+\cdots+\theta_q^2), & 0<k\le q \\ 0, & k>q \end{cases} \tag{6.17}$$

可见，当 $k>q$ 时，$\rho_k=0$，即 y_t 与 y_{t+k} 不相关，这种现象称为截尾。因此，可以根据自相关系数是否从某一点开始一直为 0 来判断 MA（q）模型的阶数，即 q 值。

MA（q）模型的偏自相关函数的具体形式随着 q 的增加变得越来越复杂，很难给出一个关于 q 的表达式，但是，一个 MA（q）模型对应一个 AR（∞）模型。由后面将要介绍的 AR（p）模型知识可知，MA（q）模型的偏自相关函数呈现出指数或者正弦形式衰减的拖尾特征。

比如一个 MA（2）模型 $y_t = 2 + u_t + 0.6u_{t-1} - 0.3u_{t-2}$ 的序列折线图如图 6—1 所示，相关分析图如图 6—2 所示。可以看出该 MA（2）模型的自相关函数从 q=3 开始截尾，偏自相关函数呈现出交替指数形式衰减的拖尾特征。

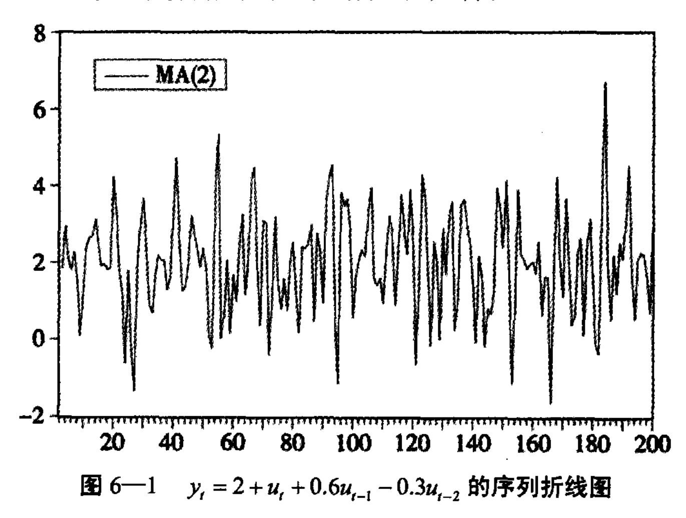


图 6—1 $y_t = 2 + u_t + 0.6u_{t-1} - 0.3u_{t-2}$ 的序列折线图

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.204	0.204	8.3506	0.004
		2	-0.363	-0.422	35.017	0.000
		3	-0.096	0.121	36.906	0.000
		4	-0.012	-0.218	36.934	0.000
		5	0.010	0.103	36.955	0.000
		6	-0.056	-0.214	37.597	0.000
		7	0.029	0.196	37.777	0.000
		8	0.135	-0.062	41.563	0.000
		9	-0.015	0.053	41.611	0.000
		10	-0.150	-0.170	46.359	0.000
		11	-0.134	-0.039	50.174	0.000
		12	0.062	0.030	50.984	0.000
		13	0.157	0.052	56.258	0.000
		14	-0.065	-0.130	57.160	0.000
		15	-0.124	0.002	60.477	0.000

图 6—2 $y_t = 2 + u_t + 0.6u_{t-1} - 0.3u_{t-2}$ 的相关分析图

（2）AR（p）模型的自相关函数和偏自相关函数

AR（p）模型

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \delta + u_t \tag{6.18}$$

的自协方差 γ_k 为：

$$\gamma_k = E(y_{t+k} y_t) = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \tag{6.19}$$

进而得到自相关函数为：

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \cdots + \phi_p\rho_{k-p} \tag{6.20}$$

可见，AR（p）模型的自相关函数是非截尾的，称为拖尾。但是 AR（p）模型的自相关函数随着滞后期 k 的增加，呈指数或者正弦形式的衰减。

由（6.13），利用 $\rho_k = \rho_{-k}$，得到如下方程组：

$$\begin{aligned}
&\rho_1 = \phi_1 + \phi_2\rho_1 + \cdots + \phi_p\rho_{p-1} \\
&\rho_2 = \phi_1\rho_1 + \phi_2 + \cdots + \phi_p\rho_{p-2} \\
&\vdots \\
&\rho_p = \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \cdots + \phi_p\rho_{p-k}
\end{aligned} \tag{6.21}$$

此方程组被称为 Yule Walker 方程组。对于 p（p=1，2，…，k，…）阶方程组求解，每个方程组的最后一个解就是相应的偏自相关系数 $\varphi_{1,1}$，$\varphi_{2,2}$，…，$\varphi_{k,k}$，…。对于 AR（p）模型，$\varphi_{k,k}$ 的最高阶数为 p，因此 AR（p）模型的偏自相关系数是 p 阶截尾的。因此，可以根据偏自相关系数的个数，来确定 AR（p）模型的阶数，即 p 值。

比如一个 AR（1）模型 $y_t = 0.6y_{t-1} + 2 + u_t$ 的序列折线图如图 6—3 所示，相关分析图如图 6—4 所示。可以看出该 AR（1）模型的偏自相关函数从 p=2 开始截尾，自相关函数呈现出指数形式衰减的拖尾特征。

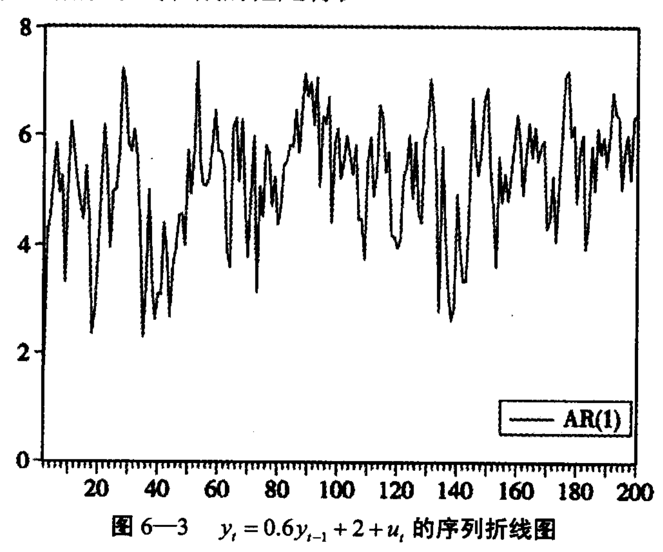


图 6—3 $y_t = 0.6y_{t-1} + 2 + u_t$ 的序列折线图

（3）ARMA（p，q）模型的自相关函数和偏自相关函数

ARMA（p，q）模型当 p，q 都不为 0 时，自相关函数和偏自相关函数都具有拖尾的特征。至于 p 和 q 的识别，则要从低阶开始逐步试探，直到定出合适的模型为止。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.530	0.530	56.790	0.000
		2	0.299	0.025	74.935	0.000
		3	0.225	0.080	85.299	0.000
		4	0.187	0.047	92.493	0.000
		5	0.196	0.087	100.39	0.000
		6	0.102	-0.075	102.55	0.000
		7	-0.002	-0.085	102.55	0.000
		8	-0.062	-0.067	103.36	0.000
		9	0.016	0.104	103.42	0.000
		10	-0.010	-0.063	103.44	0.000
		11	-0.054	-0.033	104.07	0.000
		12	-0.113	-0.075	106.81	0.000
		13	-0.094	0.032	108.69	0.000
		14	-0.019	0.045	108.77	0.000
		15	0.072	0.113	109.89	0.000

图 6—4 $y_t=0.6y_{t-1}+2+u_t$ 的相关分析图

比如一个 ARMA（1，1）模型 $y_t=0.6y_{t-1}+2+u_t+0.6u_{t-1}$ 的序列折线图如图 6—5 所示，相关分析图如图 6—6 所示。可以看出该 ARMA (1,1)模型的自相关函数呈现出指数形式衰减的拖尾特征，偏自相关函数呈现出交替指数形式衰减的拖尾特征。

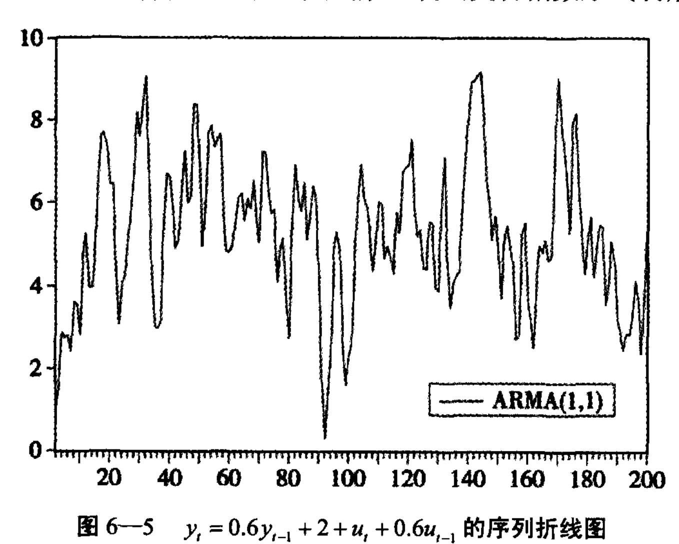


图 6—5 $y_t=0.6y_{t-1}+2+u_t+0.6u_{t-1}$ 的序列折线图

3. 线性随机时间序列模型的估计

确定模型的阶数以后，接着就可以进行参数估计。ARMA（p，q）模型的估计方法，Eviews 软件采用的是非线性方法，具体方法不作介绍，可参见时间序列分析的教科书。但要注意，MA 模型的参数估计相对困难，应尽量避免使用高阶的移动平均模型或含有高阶移动平均项的 ARMA 模型。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.782	0.782	123.45	0.000
		2	0.488	-0.317	171.75	0.000
		3	0.320	0.187	192.58	0.000
		4	0.212	-0.099	201.80	0.000
		5	0.144	0.065	206.05	0.000
		6	0.102	-0.024	208.22	0.000
		7	0.079	0.031	209.53	0.000
		8	0.044	-0.071	209.93	0.000
		9	-0.008	-0.031	209.94	0.000
		10	-0.028	0.038	210.10	0.000
		11	-0.023	-0.008	210.21	0.000
		12	-0.033	-0.043	210.44	0.000
		13	-0.061	-0.042	211.24	0.000
		14	-0.072	0.024	212.36	0.000
		15	-0.089	-0.085	214.09	0.000

图 6—6 $y_t = 0.6y_{t-1} + 2 + u_t + 0.6u_{t-1}$ 的相关分析图

4. 线性随机时间序列模型的诊断和检验

参数估计后，应该对 ARMA 模型的适合性进行检验，主要是三方面检验：

（1）检验模型参数显著性水平，即 t 统计量检验。

（2）为保证 ARMA（p，q）模型的平稳性和可逆性，模型的特征根皆应在单位圆以外，即特征根的倒数在单位圆以内。

（3）对残差序列进行白噪声检验。若序列不是白噪声，意味着残差序列还存在有用的信息没有被提取，需要进一步改进模型。残差序列的白噪声检验，可以使用上个实验介绍的 Q 统计量检验方法进行，这里不再赘述。

5. 线性随机时间序列模型的应用

（1）研究时间序列本身的变化规律，并进行拟合和预测。

（2）在回归模型的预测中首先预测解释变量的值。在我们利用结构模型来解释序列 y_t 的变动时，即使我们能够估计出一个有关 y_t 的显著性令人满意的回归方程，但其结果可能也不能用于预测。这是因为在用回归方程预测 y_t 之前，必须先得到非滞后解释变量的预测值，而这可能比预测 y_t 更困难。此时，我们可以利用 ARMA 模型来预测非滞后解释变量，作为预测 y_t 的依据。

（3）与回归模型相结合，解决残差项的序列相关问题。涉及时间序列的回归模型，残差项的序列相关问题是很常见的。通过 DW 检验和 LM 检验可以判断出残差项的序列相关问题，我们可以运用 ARMA 模型对此问题进行修正。比如回归模型：

$$y_t = X_t'\beta + u_t \tag{6.22}$$

其中：u_t 是随机扰动项，称为非条件残差。如果 u_t 存在序列相关，我们可以利用 ARMA 模型来解释 u_t，即：

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \cdots + \phi_p u_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{6.23}$$

使残差的随机扰动项成为白噪声，从而提取回归模型没有解释的那部分信息。(6.23)式的形式、定阶、参数估计、检验及预测问题都可以利用前面介绍的 ARMA 模型方法，只不过处理对象为回归模型的残差序列 u_t。

将回归模型与 ARMA 模型相结合，在 Eviews 中的实现是很方便的。比如检验发现回归模型的残差项为 AR(2)，只要在方程定义对话框中输入“y c x_1 x_2 … ar(1) ar（2）”即可。其中，ar（1）ar（2）即表示对原回归模型的残差项使用 AR（2）模型拟合。

【实验数据】

美元对英镑汇率 1990 年 1 月到 2007 年 12 月的月均价数据（参见资料光盘：数据集/ARMA 模型应用数据/美元对英镑汇率月均价数据.xls）。数据来源为 The University of British Columbia Sauder School of Business Pacific Exchange Rate Service，网址为：http://fx.sauder.ubc.ca/。

【实验内容】

外汇汇率是不同货币之间兑换的比率或比价，是以一种货币表示的另一种货币的价格。造成汇率波动的因素很多，例如，国际收支差额、利率水平、通货膨胀、财政政策和货币政策、投资资本、政府的市场干预、一国经济实力和其他非市场因素。反过来，汇率波动也是对上述因素的一个反映。本次实验，同学们可以根据美元对英镑汇率 1990 年 1 月到 2007 年 12 月的月均价数据，利用 ARIMA（p，d，q）模型对该序列的变化特征和规律进行分析，同时选择最优的 ARIMA（p，d，q）模

型形式对 2008 年 1 月美元对英镑的汇率进行预测。

【实验步骤】

1. 根据数据频率和时间范围，创建 Eviews 工作文件（Workfile）。
2. 录入数据，并对序列进行初步分析。绘制美元对英镑汇率的月均价数据序列的折线图，分析序列的基本趋势，初步判断序列的平稳性。
3. 识别 ARIMA（p，d，q）模型中的阶数 p，d，q。运用单位根检验（ADF 检验）确定单整阶数 d；利用相关分析图来确定自回归阶数 p 和移动平均阶数 q。初步选择几个适合的备选模型。
4. ARIMA（p，d，q）模型的估计与检验。对备选模型进行估计和检验，并进行比较，从中选择最优模型。
5. 利用择最优模型对 2008 年 1 月美元对英镑汇率的月均价进行外推预测。
6. 综合上述实验步骤得出的结果，得出最终结论。总结实验过程中的问题以及得到的经验教训，完成实验报告。

【问题思考】

1. 什么样的序列可以应用 ARIMA（p，d，q）模型？
2. ARIMA（p，d，q）模型中的阶数 p，d，q 如何识别？
3. ARIMA（p，d，q）模型的应用主要有哪几个方面？

【实验总结】

通过本实验，加深对 ARIMA（p，d，q）模型的各种形式和基本原理的理解和掌握。熟练掌握 ARIMA（p，d，q）模型中的阶数 p，d，q 的识别方法。学会 ARIMA（p，d，q）模型的建立、检验及预测的基本方法。写出实验报告。

实验七　条件异方差模型应用

【实验目的与要求】

1. 准确掌握条件异方差模型的各种形式和基本原理。
2. 熟练掌握检验 ARCH 效应的方法。
3. 学会 GARCH（1，1）模型和 GARCH-M 模型的建立及检验方法。
4. 熟练掌握运用回归-ARCH 模型对样本序列进行拟合和预测。
5. 在教师的指导下独立完成实验，得出正确的结果，并完成实验报告。

【实验准备知识】

在金融时间序列的分析中，经常会遇到观测值在某个时间段变化幅度大，在另一个时间段变化幅度又比较小的情况。相应的，用模型预测的时候，误差在某一时期里较大，而在另一时期里相对较小。这种性质被称为聚类性（clustering）。这类序列随机扰动项的无条件方差是常量，而条件方差是变化的量，为了描述这类序列的特征，我们引入所谓的条件异方差模型。

1. ARCH 模型

Engle 于 1982 年提出的自回归条件异方差模型（Auto-Regressive Conditional Heteroskedasticity model，ARCH 模型）是最基本的模型。它的基本思想是：扰动项 u_t 的条件方差依赖于它的前期值的大小。对于通常的回归模型：

$$y_t = X\beta + u_t \tag{7.1}$$

如果随机扰动项的平方u_t^2服从AR(q)过程，即：

$$u_t^2=\alpha_0+\alpha_1 u_{t-1}^2+\alpha_2 u_{t-2}^2+\cdots+\alpha_q u_{t-q}^2+\varepsilon_t \tag{7.2}$$

其中：ε_t是白噪声，若满足ε_t~IID（0，λ^2）则称上述模型是自回归条件异方差模型，简记为ARCH模型，称序列u_t服从q阶的ARCH的过程，记作u_t~ARCH（q）。模型还可以写成：

$$u_t=\sqrt{h_t}v_t \tag{7.3}$$

$$h_t=\alpha_0+\alpha_1 u_{t-1}^2+\cdots+\alpha_q u_{t-q}^2=\alpha_0+\sum_{i=1}^{q}\alpha_i u_{t-i}^2 \tag{7.4}$$

其中：v_t~IID（0，1）。对于任意时刻t，u_t的条件期望为0，条件方差为h_t。为了保证h_t为正值，要求$\alpha_0>0$，$\alpha_i\geqslant 0$（i=1,2,3,4，…，q）。同时，为了保证ARCH过程的平稳性，要求$\sum_{i=1}^{q}\alpha_i<1$。

由于y_t的均值近似等于（7.1）式的估计值，所以（7.1）式也称为均值方程，而（7.2）式或（7.4）式也称为条件方差方程。

从ARCH模型中可以看出，由于当期随机扰动项的方差是过去有限项随机扰动项值平方的回归，也就是说随机扰动项的波动具有一定的记忆性，因此，如果前期随机扰动项的方差较大，那么当期随机扰动项的方差往往也较大，反之亦然。自回归阶数q决定了冲击的影响存留于后续随机扰动项方差中的时间长度，q值越大，波动持续的时间也就越长。因此，ARCH模型具有描述波动的聚类性的能力。

（7.1）式和（7.2）式构成的模型被称为回归–ARCH模型。ARCH模型通常用于对主体模型的随机扰动项进行建模分析。从而充分提取残差中的信息，使得最终的模型残差ε_t成为白噪声序列。

2. GARCH模型

许多经济问题常常出现u_t的条件方差h_t依赖于很多时刻之前的变化量的现象，这样如果使用ARCH模型，阶数q需要取一个很大的值。为了避免估计很多个参数，我们考虑使用广义自回归条件异方差模型（Generalized Auto-Regressive Conditional

Heteroskedasticicity model，GARCH 模型），该模型由 Bollerslev 于 1986 年提出。相对于 ARCH，GARCH 模型的优点在于：可以用较为简单的 GARCH 模型来代表一个高阶 ARCH 模型，从而使得模型的识别和估计都变得比较容易。

若（7.4）式为下面的形式：

$$\begin{aligned} h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_q u_{t-q}^2 + \beta_1 h_{t-1} + \cdots + \beta_p h_{t-p} \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \end{aligned} \tag{7.5}$$

则称序列服从 GARCH（p，q）过程。其中：$p \geqslant 0$；$q>0$；$\alpha_0>0$；$\alpha_i \geqslant 0$（i=1,2,3,4，⋯，q）；$\beta_j \geqslant 0$（j=1,2,3,4，⋯，p）。可见，ARCH（q）过程只是 GARCH（p，q）过程的一个特例。

同时，为了保证 GARCH 过程的平稳性，要求$\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1$。从另一个角度讲，可以说明外部冲击对干扰项$u_t$的波动特征产生的影响将随时间的推移而逐渐衰减。以 GARCH（1，1）模型为例，即$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}$，两边取期望可得：

$$E(h_t) = \alpha_0 + (\alpha_1 + \beta_1) h_{t-1} \tag{7.6}$$

通过迭代，可得：

$$\begin{aligned} E(h_{t+1}) &= \alpha_0 + (\alpha_1 + \beta_1) h_t \\ E(h_{t+2}) &= \alpha_0 + (\alpha_1 + \beta_1) h_{t+1} = \alpha_0 + (\alpha_1 + \beta_1)[\alpha_0 + (\alpha_1 + \beta_1) h_t] \\ &= \alpha_0 [1 + (\alpha_1 + \beta_1)] + (\alpha_1 + \beta_1)^2 h_t \\ &\cdots \\ E(h_{t+m}) &= \alpha_0 [1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 + \cdots + (\alpha_1 + \beta_1)^{m-1}] + (\alpha_1 + \beta_1)^m h_t \end{aligned} \tag{7.7}$$

从（7.7）式可以看出，当 t 时刻某外部冲击使条件方差 h_t 发生变化时，t 后一段时间内条件方差都会受到影响，其大小取决于$(\alpha_1 + \beta_1)^m$。$(\alpha_1 + \beta_1) < 1$说明该冲击对未来各时刻的影响将呈指数衰减，$(\alpha_1 + \beta_1)$称为衰减系数。$(\alpha_1 + \beta_1)$越大，衰减的速度越慢。

3. ARCH-M 模型和 GARCH-M 模型

一般来说，每个投资者都认为资产（尤其是金融资产）的收益率与其风险正相关，风险越大，预期的收益就越高。因此，为了更好地将收益与风险联系起来，将

代表预期风险的条件方差 h_t（或者标准差 $\sqrt{h_t}$）作为影响序列 y_t 本身的解释变量之一，引入 ARCH 模型的均值方程中，这种模型被称为 ARCH-M(ARCH-in-Mean)模型。其表达式为：

$$y_t = X\beta + \gamma h_t + u_t \tag{7.8}$$

$$u_t = \sqrt{h_t} v_t \tag{7.9}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 \tag{7.10}$$

其中，参数 γ 度量了条件方差 h_t 对 y_t 的影响程度，它代表了风险和收益之间的一种权衡。

如果条件方差 h_t 的结构为 GARCH（p，q）过程，即：

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{7.11}$$

则模型称为 GARCH-M 模型。GARCH-M 模型通常应用于资产的预期收益与预期风险密切相关的金融领域。

4. ARCH 效应检验

一个模型的残差是否含有 ARCH 效应，可以通过以下几种方法检验。

（1）ARCH-LM 检验

检验 ARCH 效应最常用的方法是拉格朗日乘数法检验，即 ARCH-LM 检验。建立辅助回归方程

$$\hat{u}_t^2 = \alpha_0 + \alpha_1 \hat{u}_{t-1}^2 + \cdots + \alpha_q \hat{u}_{t-q}^2 + \varepsilon_t \tag{7.12}$$

其中：$\hat{u}_t$ 是残差。

检验残差序列是否存在 ARCH 效应，即检验（7.12）式中的回归系数是否同时为 0。检验的原假设是：

$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$（即不存在 ARCH 效应）

检验统计量：

$$LM = TR^2 \sim \chi^2(q) \tag{7.13}$$

其中：T 是计算辅助回归（7.12）式的样本数据个数；R^2 是（7.12）式的样本决定系数。

判别规则是：给定显著性水平 α 和自由度 q，若 $LM > \chi_\alpha^2(q)$，则拒绝 H_0，认为残差序列存在 ARCH 效应；

若 $LM \le \chi_\alpha^2(q)$，则接受 H_0，认为残差序列不存在 ARCH 效应。

（2）残差平方的 Q 统计量检验

在实验四中，我们介绍过 Q 统计量检验，Q 统计量的形式为：

$$Q = T(T+2)\sum_{k=1}^{K}\frac{r_k^2}{T-k} \tag{7.14}$$

如果估计的自相关系数 r_k 是用残差值的平方序列计算的，那么 Q 统计量就可以检验残差序列中是否存在自回归条件异方差（ARCH、GARCH 过程）。这时的零假设是残差序列不存在 ARCH 效应；备择假设是残差序列存在 ARCH 效应。检验方法与所用临界值与实验四中介绍的 Q 统计量检验相同。

（3）残差平方序列的相关分析图

残差平方序列的相关分析图显示残差平方序列的自相关系数和偏自相关系数。如果残差序列不存在 ARCH 效应，残差平方序列的自相关系数和偏自相关系数在所有的滞后阶数都应为零，否则，就说明残差序列存在 ARCH 效应。

【实验数据】

我国上海证券交易所 B 股指数 1998 年 1 月 9 日到 2008 年 3 月 7 日周收盘价数据（参见资料光盘：数据集/条件异方差模型应用数据/上证 B 股指数周数据.xls）。数据来源为大智慧软件（下载并整理）。

【实验内容】

股票价格指数，是反映股票市场总体行情的重要指标。股票价格指数序列常常可以用随机游走模型来描述，但是模型的残差通常会出现 “聚类现象”，即所谓的条件异方差性（ARCH 效应）。为了描述和消除这种效应，我们通常使用条件异方差模型。本次实验，同学们可以根据我国上海证券交易所 B 股指数 1998 年 1 月 9 日到 2008 年 3 月 7 日周收盘价数据，利用 GARCH(1,1)模型和 GARCH-M 模型建立回归–ARCH 模型，对该序列进行拟合。同时选择最优的条件异方差模型形式对 2008 年 3 月 14 日上证 B 股指数收盘价进行外推预测。

【实验步骤】

1. 根据数据频率和时间范围，创建 Eviews 工作文件（Workfile）。

2. 录入数据，并对序列进行初步分析。分别绘制 B 股指数周收盘价数据原序列和对数序列的折线图，初步分析序列的基本趋势和波动特征。

3. 建立主体模型。用对数序列建立一阶自回归模型作为主体模型。观察残差序列图和残差平方序列图，初步判断 ARCH 效应。

4. ARCH 效应检验。利用 ARCH-LM 检验、残差平方序列的相关分析图和残差平方的 Q 统计量检验，对一阶自回归模型残差的 ARCH 效应进行检验。

5. 建立条件异方差模型。利用 GARCH(1,1)模型和 GARCH-M 模型分别建立回归–ARCH 模型，并进行比较，从中选择最优模型。

6. 利用择最优模型对 2008 年 3 月 14 日上证 B 股指数收盘价进行外推预测。

7. 综合上述实验步骤得出的结果，得出最终结论。总结实验过程中的问题以及得到的经验教训，完成实验报告。

【问题思考】

1. 什么是 ARCH 效应？如何识别？
2. 条件异方差模型如何解决残差的 ARCH 效应问题？
3. GARCH-M 模型有几种形式？相对于 GARCH 模型有什么优点？

【实验总结】

通过本实验，加深对条件异方差模型的各种形式和基本原理的理解和掌握。熟练掌握 ARCH 效应的识别方法。学会 GARCH 模型和 GARCH-M 模型的建立、检验及预测的方法。写出实验报告。


案例篇

案例一　招商银行A股β系数估计

在本案例中，我们利用我国上海证券交易所上证指数（代码 000001）和招商银行 A 股（代码 600036）2002 年 4 月到 2008 年 2 月的月收盘价数据[①]（参见资料光盘：数据集/线性回归模型应用数据/上证指数月数据.xls 和招商银行月数据.xls），介绍 Eviews 软件的一些基本操作以及如何建立、估计、检验线性回归模型。

1. 创建 Eviews 工作文件（Workfile）

Eviews 软件是一款基于 Windows 操作系统的计量分析软件，它在数据分析与评价、金融分析、经济预测、销售预测和成本分析等领域有着非常广泛应用。本书的实验以及案例分析都是应用 Eviews 软件进行的。

Eviews 软件要求数据的处理和分析必须在特定的工作文件（Workfile）中进行，因此我们首先建立工作文件。在正确安装 Eviews 软件的基础上，启动 Eviews 软件，出现 Eviews 主界面，如图 1—1 所示。

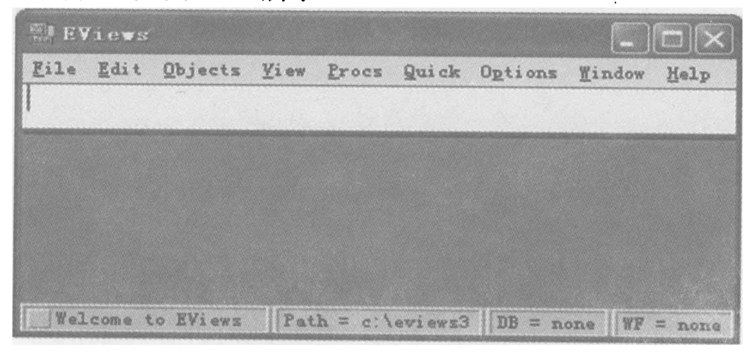


图 1—1　Eviews 主界面

① 数据来源：大智慧软件。

建立工作文件的方法是：从Eviews主选单中选“File/New/Workfile”，如图1—2所示。

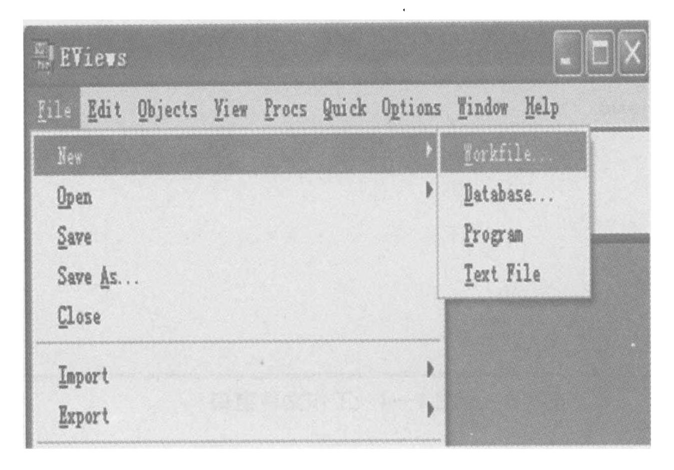


图1—2　建立工作文件

点击后，出现如图1—3所示的数据频率及范围选择对话框。

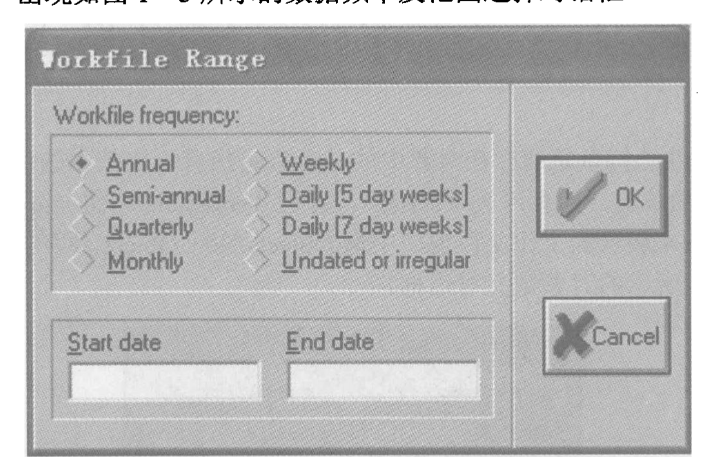


图1—3　数据频率及范围选择

Annual（年度数据）：可以用四位年份，如“Start date：1955”、“End date：2007”，在1900年和2000年之间的年份只需要输入后两位即可。

Quarterly（季度数据）：输入格式为“1992:1”和“2006:4”，年后面只能跟1、2、3、4代表季度。

Monthly（月度数据）：输入格式为“1956:1”和“2006:11”，年后面为月。

Weekly and daily（周或日数据）：输入格式为“月:日:年”，如“8:10:2007”，即2007年8月10日。

Undated or irregular（非日期数据）：如“Start date：1”、“End date500”，即500个数的一个序列。

本例中，选择Monthly选项，在Start date后输入“2002:04”，在End date后“2008:

02”，再点击“OK”，出现如图1—4所示的工作文件窗口。

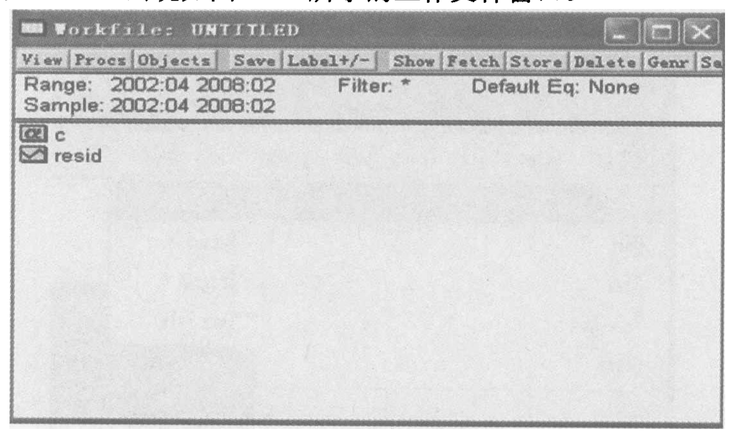


图1—4 工作文件窗口

2. 录入数据，并对序列进行初步分析

Eviews软件的信息是储存在对象中的，对象是指有一定关系的信息或算子捆绑在一起供使用的单元，包括序列（series）、方程（equation）、组（group）等。

在如图1—4所示的工作文件窗口中选Objects/New Object，即可新建一个对象，出现如图1—5所示的对象定义对话框。

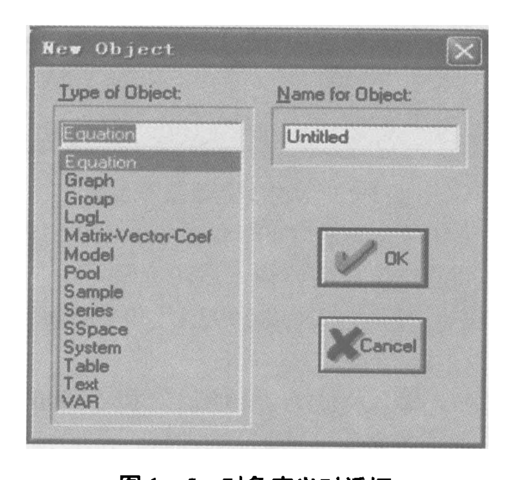


图1—5 对象定义对话框

我们选择“Series（序列对象）”，命名为sh，用来保存上证指数月收盘价数据，点击“OK”后会发现工作文件中多了一个sh序列对象，双击该对象打开，如图1—6所示。可以看到没有导入数据之前，该对象相当于一个一列的空白表格。

Series: SH　Workfile: UNTITLED

View | Procs | Objects | Print | Name | Freeze | Edit+/- | Smpl+/- | Label+/- | Wi

	SH
	Last updated: 05/07/08 - 11:17
2002:04	NA
2002:05	NA
2002:06	NA
2002:07	NA
2002:08	NA
2002:09	NA
2002:10	NA
2002:11	NA
2002:12	NA
2003:01	

图 1—6　sh 空白序列

点击“Edit+/-”，允许对数据进行编辑，本例中，我们将 Excel 文件中的数据拷贝进来即可，结果如图 1—7 所示。

Series: SH　Workfile: UNTITLED

View | Procs | Objects | Print | Name | Freeze | Edit+/- | Smpl+/- | Label+/- | Wi

1667.749023　SH

	Last updated: 05/07/08 - 11:29
2002:04	1667.749
2002:05	1515.733
2002:06	1732.755
2002:07	1651.594
2002:08	1666.618
2002:09	1581.618
2002:10	1507.496
2002:11	1434.182
2002:12	1357.654
2003:01	

图 1—7　导入数据

单从数字上我们很难出序列的特征，因此需要观察图形。在序列窗口下选中“View/Line Graph”，如图 1—8 所示。

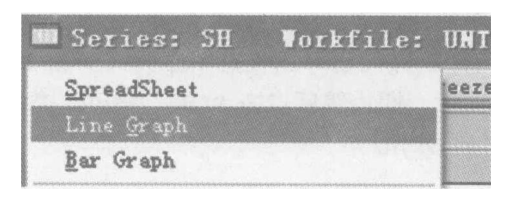


图 1—8　绘制折线图

点击后，得到上证指数月收盘价数据的折线图，如图 1—9 所示。

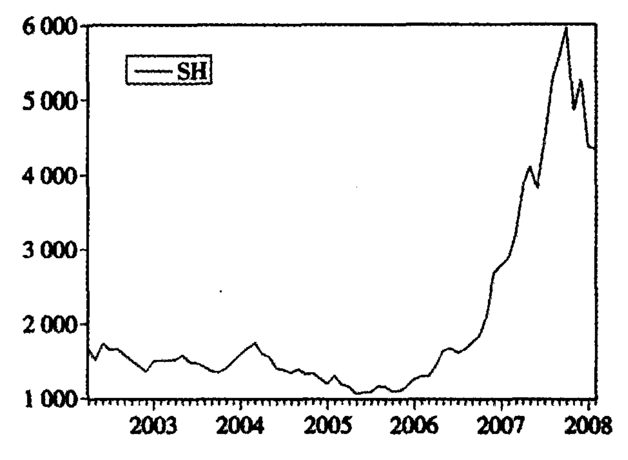


图 1—9 上证指数月收盘价数据的折线图

同样的方法，建立 zsyh 序列对象，保存招商银行 A 股月收盘价数据，并绘制折线图，如图 1—10 所示。

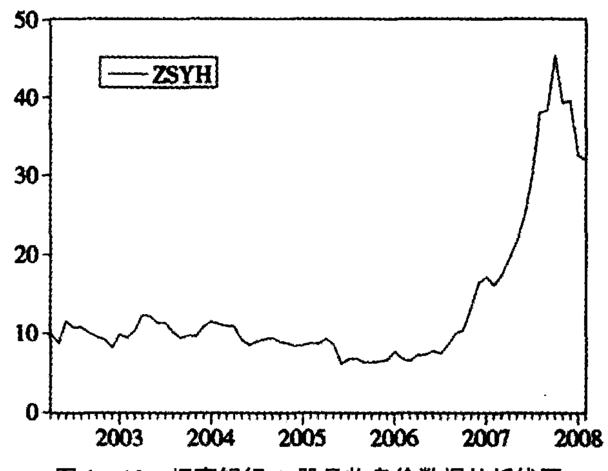


图 1—10 招商银行 A 股月收盘价数据的折线图

从图 1—9 和图 1—10 可以看出，招商银行 A 股和上证指数的走势基本一致。为了利用单指数模型估计招商银行 A 股的 β 系数，我们还需要建立招商银行 A 股和上证指数的收益率序列，这里我们取百分比收益率。

从 Eviews 主选单中选“Quick/Generat Series”，如图 1—11 所示。

点击后，出现序列对象定义对话框，输入“rsh=（sh-sh（－1））/sh（－1）”，建立一个新序列 rsh，为上证指数的百分比收益率序列（如图 1—12 所示）。其中，sh（－1）表示 sh 序列的滞后一期。然后点击“OK”即可。类似的方法，建立招商银行 A 股百分比收益率序列 rzsyh。

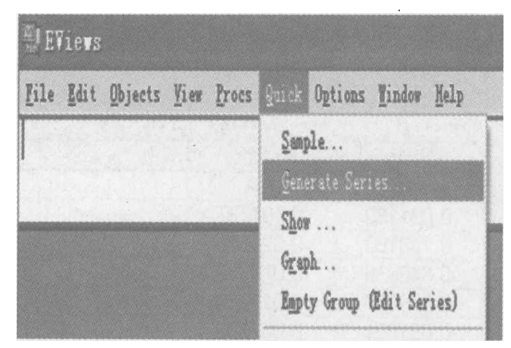


图1—11　用公式的形式创建序列

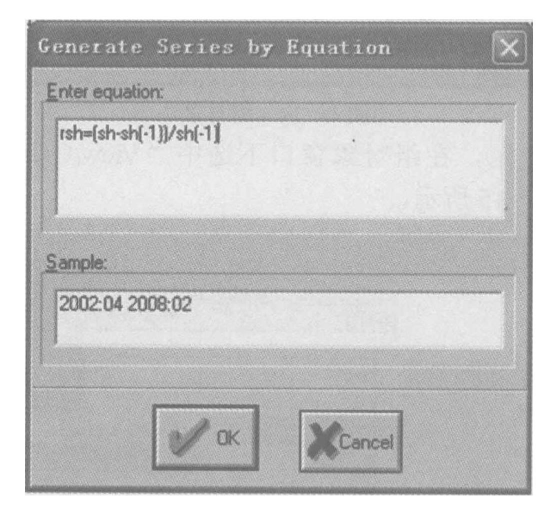


图1—12　建立上证指数的百分比收益率序列

为了确定招商银行A股和上证指数收益的相关关系，可以将rsh和rzsyh序列以组的形式进行观察。在工作文件界面，同时选择rsh和rzsyh序列对象，方法是按住键盘Ctrl键，鼠标左键点击两个对象图标，再点击鼠标右键，选择“Open/as Group”，即以组对象的形式打开，如图1—13所示。

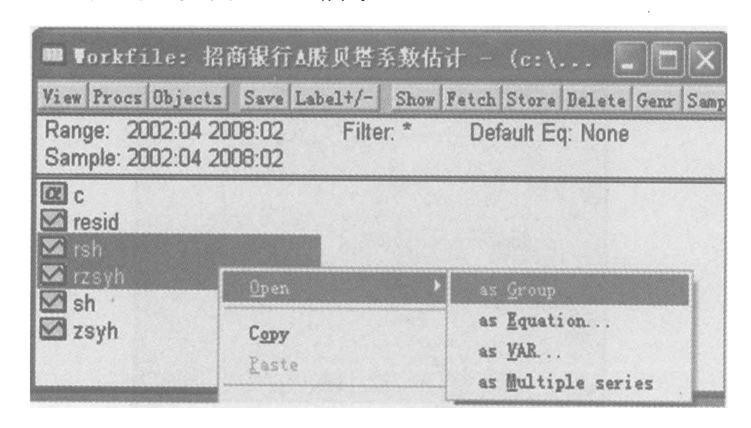


图1—13　以组对象的形式打开rsh和rzsyh序列对象

点击后，结果如图 1—14 所示。

Group: UNTITLED　Workfile: 招商银行A...

View | Procs | Objects | Print | Name | Freeze | Edit+/- | Smpl+/- | InsDel | Trans

obs	RSH	RZSYH		
2002:04	NA	NA		
2002:05	-0.091150	-0.131737		
2002:06	0.143180	0.327586		
2002:07	-0.046839	-0.073593		
2002:08	0.009097	0.010280		
2002:09	-0.051001	-0.067530		
2002:10	-0.046865	-0.050595		
2002:11	-0.048633	-0.037618		
2002:12	-0.053360	-0.116178		
2003:01				

图 1—14　组对象

先观察二维散点图，在组对象窗口下选中“View/Graph/Scatter/Scatter with Regresstion”，如图 1—15 所示。

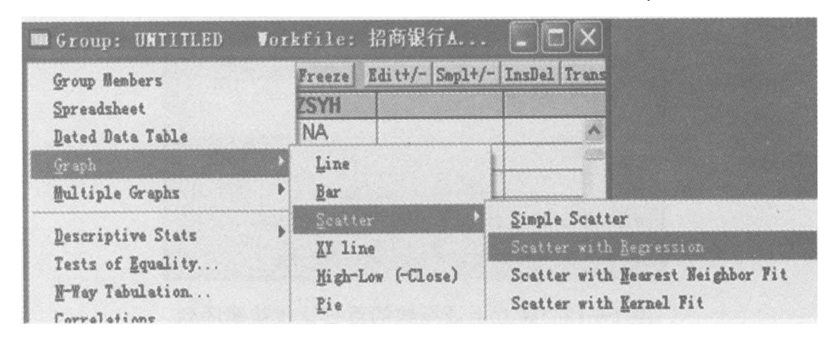


图 1—15　绘制 rsh 和 rzsyh 序列二维散点图

点击后，出现散点图绘制选项对话框，如图 1—16 所示。

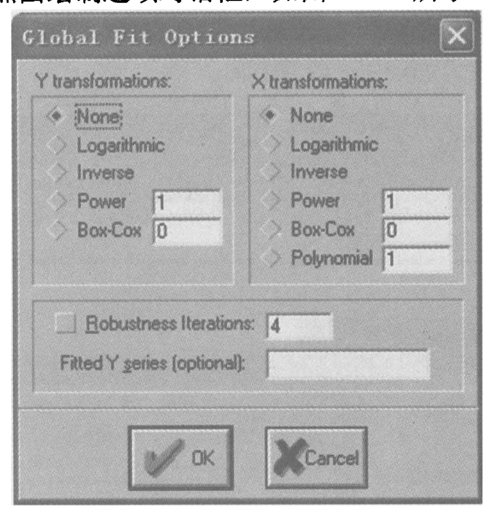


图 1—16　散点图绘制选项对话框

保持默认选项，点击"OK"，结果如图 1—17 所示。

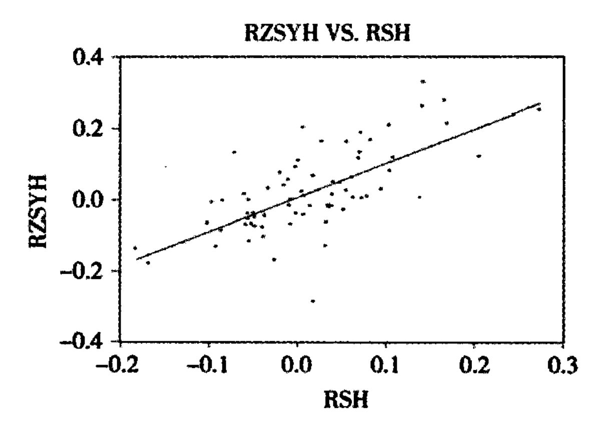


图 1—17 rsh 和 rzsyh 序列二维散点图

从图 1—17 可看出，大多数散点分布在一条直线（回归线）附近，可以初步认为 rsh 和 rzsyh 线性相关。

进一步观察 rsh 和 rzsyh 序列的相关系数，在组对象窗口下选中"View/Correlations"，如图 1—18 所示。

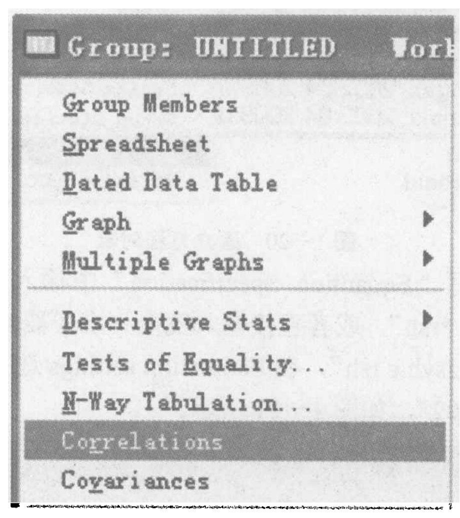


图 1—18 计算 rsh 和 rzsyh 序列的相关系数

点击后，出现 rsh 和 rzsyh 序列的相关系数矩阵，如图 1—19 所示。可见，rsh 和 rzsyh 序列的相关系数为 0.675。

Correlation Matrix		
	RSH	RZSYH
RSH	1.000000	0.675208
RZSYH	0.675208	1.000000

图 1—19 rsh 和 rzsyh 序列的相关系数矩阵

3. 建立单指数模型

（1）建立模型

为了估计招商银行 A 股的β系数，可以建立单指数模型，模型形式为：$r_{it}=a+br_{mt}+u_t$。从 Eviews 主选单中选中“Quick/Estimate Equation”，或者用新建对象的方法新建一个 Equation 对象，如图 1—20 所示。

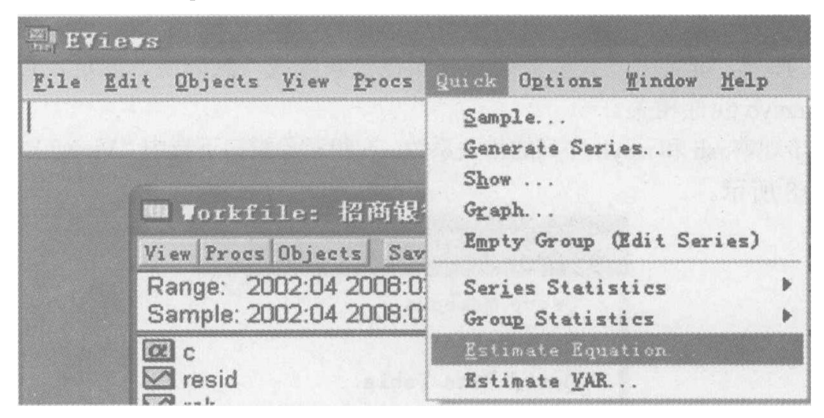


图 1—20 建立方程对象

在方程定义对话框“Equantion Specification”中输入方程的形式，可以输入“rzsyh=c（1）+c（2）*rsh”，或者直接输入变量列表（被解释变量在前，解释变量在后）。本例中输入“rzsyh c rsh”，在 Estimation settings 选择框选择 LS 估计法和样本区间“2002:04 2008:02”，如图 1—21 所示。

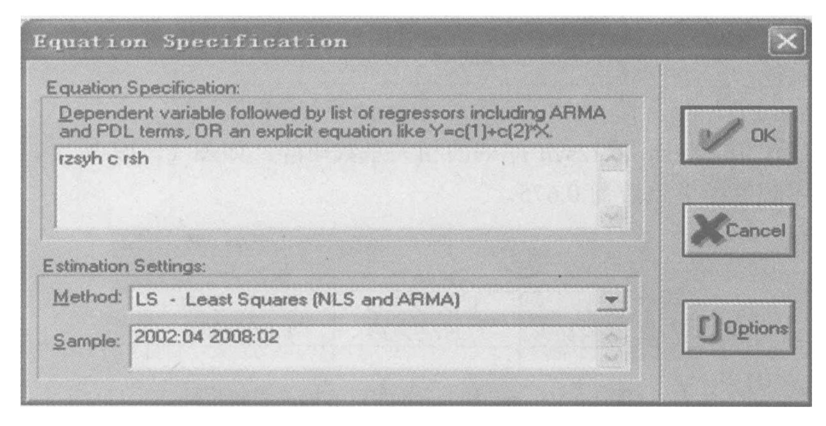


图 1—21 对方程进行定义

点击“OK”键后，得到如图1—22所示的估计结果。

Equation: UNTITLED　Workfile: 招商银行A股...

View | Procs | Objects | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: RZSYH
Method: Least Squares
Date: 05/09/08　Time: 12:43
Sample(adjusted): 2002:05 2008:02
Included observations: 70 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.006918	0.010519	0.657717	0.5129
RSH	0.953098	0.126265	7.548407	0.0000

R-squared	0.455906	Mean dependent var	0.023135
Adjusted R-squared	0.447905	S.D. dependent var	0.115943
S.E. of regression	0.086149	Akaike info criterion	-2.037322
Sum squared resid	0.504672	Schwarz criterion	-1.973079
Log likelihood	73.30626	F-statistic	56.97845
Durbin-Watson stat	2.294836	Prob(F-statistic)	0.000000

图1—22　方程输出结果

估计结果的上部显示被解释变量、估计方法、估计的时间、样本范围、样本个数；中部显示系数的估计值、标准误、*t*统计量及其相伴概率；下部显示其他一些统计量，左边是：样本决定系数、调整后的样本决定系数、回归标准误差、残差平方和、对数似然比、*DW* 统计量，右边是：被解释变量的均值、被解释变量标准差、AIC准则、SC准则、*F*检验的统计量及其相伴概率。

常数项的*t*统计量值为0.658，相伴概率为0.513，没有通过显著性检验，因此，应从方程中剔除，重新定义方程为rzsyh rsh。得到回归结果如图1—23所示（这里只截取中部和下部）。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RSH	0.970060	0.123094	7.880668	0.0000

R-squared	0.452445	Mean dependent var	0.023135
Adjusted R-squared	0.452445	S.D. dependent var	0.115943
S.E. of regression	0.085794	Akaike info criterion	-2.059552
Sum squared resid	0.507883	Schwarz criterion	-2.027430
Log likelihood	73.08431	Durbin-Watson stat	2.287109

图1—23　剔除常数项的方程输出结果

对应的回归表达式是

$$RZSYH=0.970\times RSH$$

（7.88）　　　$R^2=0.452$，$DW=2.28$

其中：括号内的值是 t 统计量。RSH 的系数估计值为 0.970，t 统计量的相伴概率为 0，通过显著性检验，也就是说招商银行 A 股的 β 系数的估计值为 0.970，跟 1 很接近，说明招商银行 A 股的价格波动基本与市场一致，这与对图 1—9 和图 1—10 的观察结果一致。R^2=0.452，说明方程的拟和程度并不是很高，但是作为金融资产的收益率序列拟合，还是可以接受的。其经济意义是，招商银行 A 股的系统风险占总风险的比例为 45.2%。DW=2.28，说明基本上没有一阶自相关。

我们选中"View/Actual，Fitted，Residual/Actual，Fitted，Residual Graph"，还可以观察真实值、拟合值与残差图，如图 1—24 所示。

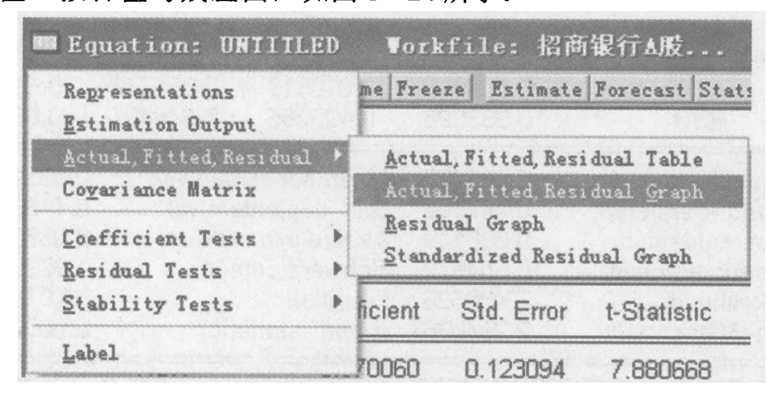


图 1—24 选择绘制真实值、拟合值与残差图

点击后，结果如图 1—25 所示。可以看到残差没有自相关，这与 DW = 2.28 说明了同样的问题。

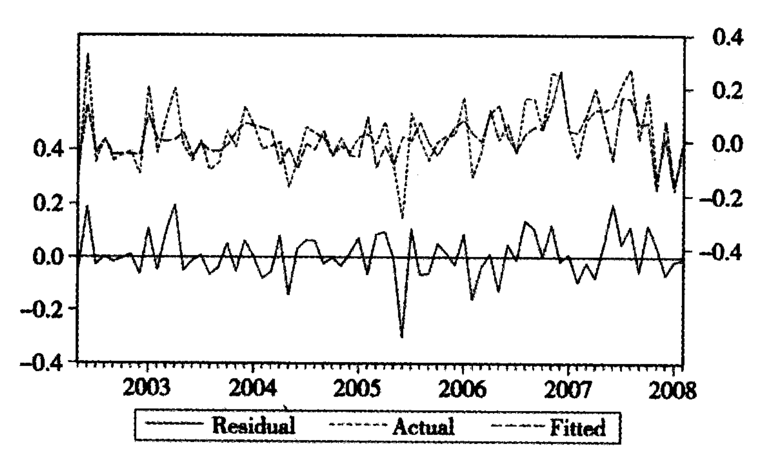


图 1—25 线性趋势模型的真实值、拟合值与残差图

（2）模型检验

除了在方程估计结果中给出的一些检验结果外，我们还可以对模型做一些检验。

①残差正态性检验。在方程窗口中选中"View/Residual Tests/Histogram–Normality Test"，如图 1—26 所示。

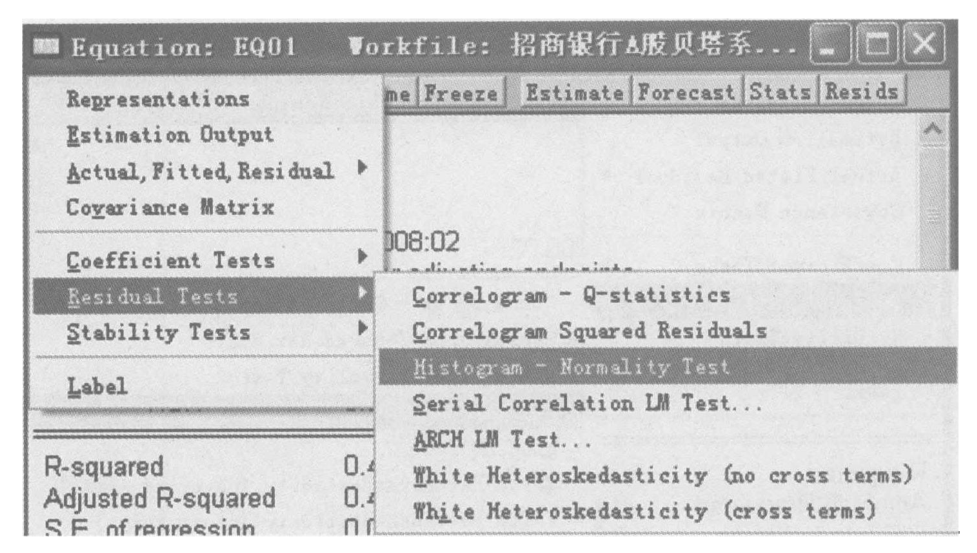


图 1—26　残差正态性检验

点击后，结果如图 1—27 所示。可以看到残差序列的偏度为−0.334，峰度为 4.586，*JB* 统计量为 8.638（四舍五入），*JB* 统计量的相伴概率为 0.013，小于 0.05，可以认为在 5%的显著性水平上认为残差序列不服从正态分布。

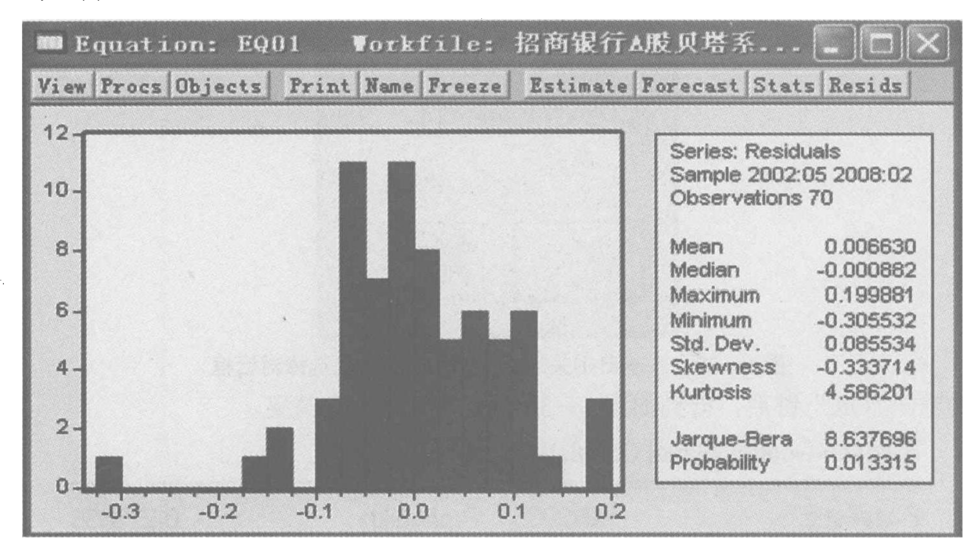


图 1—27　残差正态性检验结果

②残差序列自相关的 *LM* 检验。在方程窗口中选中“View/ Residual Tests/ Serial Correlation LM Test”，如图 1—28 所示。

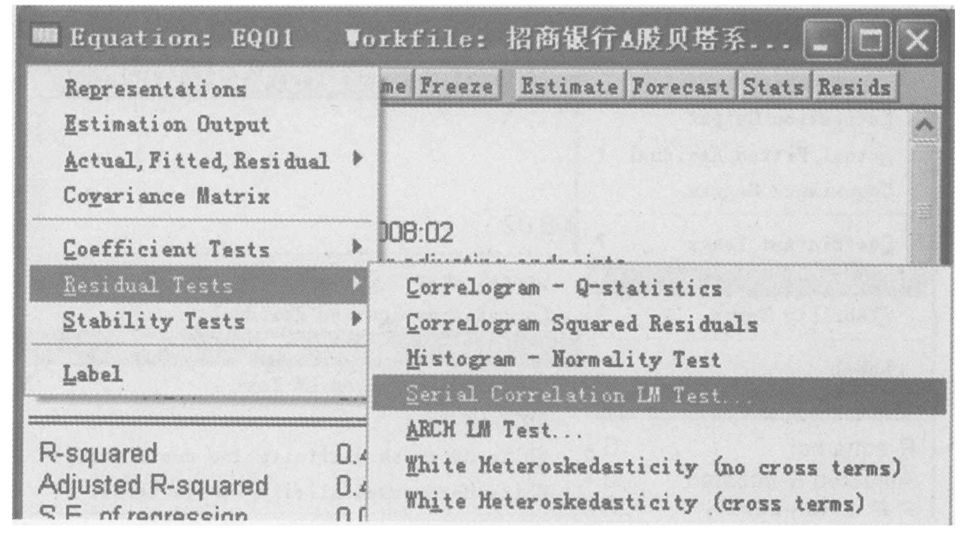


图 1—28　残差自相关的 *LM* 检验

点击后，出现滞后阶数选择对话框，如图 1—29 所示。默认值为 2，我们可以先修改为 1。



图 1—29　残差自相关的 *LM* 检验滞后阶数选择对话框

点击“OK”键后，得到如图 1—30 所示的 *LM* 检验结果。

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.512035	Probability	0.223066
Obs*R-squared	1.105306	Probability	0.293105

图 1—30　*LM* 检验结果

LM 统计量的相伴概率为 0.293，大于 0.05，可以认为在 5%的显著性水平上，不存在一阶自相关。这与 *DW* 检验结果一致。用类似的方法，我们检验出残差序列不存在二阶、三阶、四阶自相关。更进一步的高阶自相关只要输入对应的滞后阶数进行检验即可，还可以用实验四将要介绍的 *Q* 统计量进行检验。

③残差序列异方差的 White 检验。在方程窗口中选中“View/ Residual Tests/ White

Heteroskedasticity (no cross terms)”，这里选择的是没有交叉积项，也可以选择有交叉积项（cross terms），如图 1—31 所示。

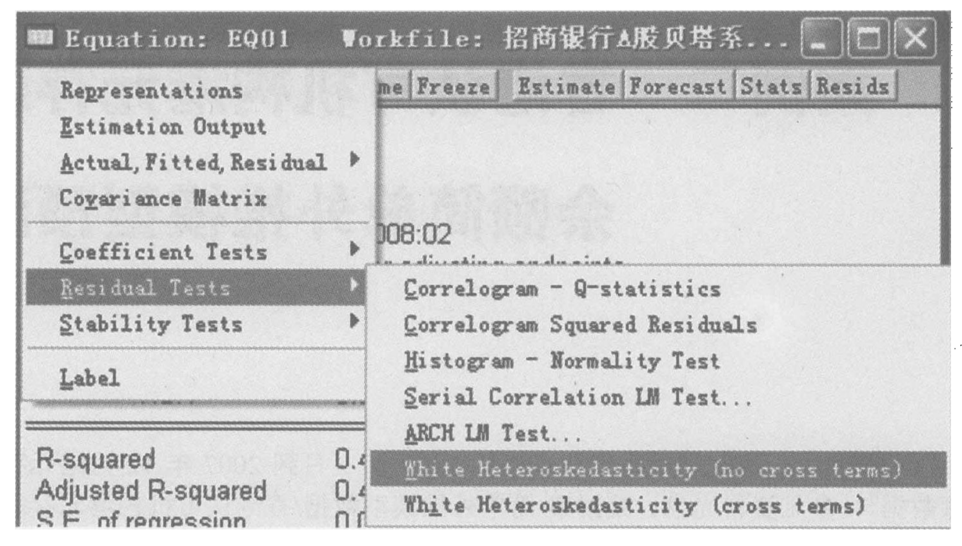


图 1—31　残差异方差的 White 检验

点击后，结果如图 1—32 所示。

White Heteroskedasticity Test:

F-statistic	0.709057	Probability	0.495763
Obs*R-squared	1.450901	Probability	0.484106

图 1—32　White 检验结果

可见，nR^2 的相伴概率为 0.484，大于 0.05，可以认为在 5%的显著性水平上，不存在异方差。

4. 小结

本例中，我们利用单指数模型对招商银行 A 股的β系数的进行估计。建立的一元线性回归方程总的效果不错：系数通过显著性检验，残差没有自相关和异方差。但不足的是正态性检验没有通过，样本决定系数 R^2=0.452 不是很高，但是作为金融资产的收益率序列拟合，还是可以接受的。

最后，得到的招商银行 A 股的β系数估计值为 0.970，跟 1 很接近，说明招商银行 A 股的价格波动基本与市场一致。同时，方程的样本决定系数 R^2=0.452，说明招商银行 A 股的系统风险占总风险的比例为 45.2%。

案例二　香港认可机构港元存款余额简单外推模型预测

在本案例中，我们利用香港认可机构 1999 年 1 月到 2007 年 12 月港元存款余额数据①（参见资料光盘：数据集/简单外推模型数据/香港认可机构港元存款余额月度数据.xls），介绍怎样利用 Eviews 软件进行简单外推模型预测，模型包括线性趋势模型、指数增长模型、自回归趋势模型、对数自回归趋势模型和二次趋势模型。

1. 创建 Eviews 工作文件（Workfile）

从 Eviews 主选单中选“File/New/Workfile”，选择“Monthly”选项，输入“Start date：1999:01 End date：2007:12”，方法如案例一介绍。

2. 录入数据，并对序列进行初步分析

在 Workfile 窗口中选“Objects/New Object”，新建一个序列对象，命名为 ck，用来保存香港认可机构港元存款余额数据，并将数据导入。该序列的折线图如图 2—1 所示。

① 数据来源　香港金融管理局：载《金融数据月报》，2008（163）。

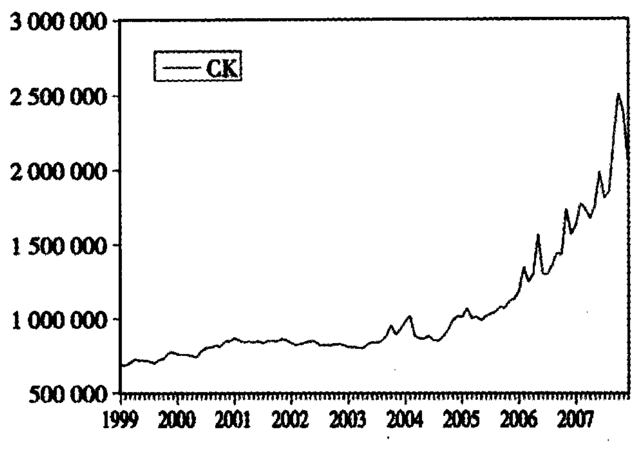


图 2—1　港元存款余额数据的折线图

从图形上我们初步看出该序列具有一个长期增长的趋势。2005 年以前增长速度慢，2005 年以后增长速度变快。

再建立时间项 t 序列。从 Eviews 主选单中选“Quick/Generat Series”，输入“t=@trend（1998:12）”，建立一个新序列，表示从 1999 年 1 月开始，该序列为 1，2，3，…，如图 2—2 所示。

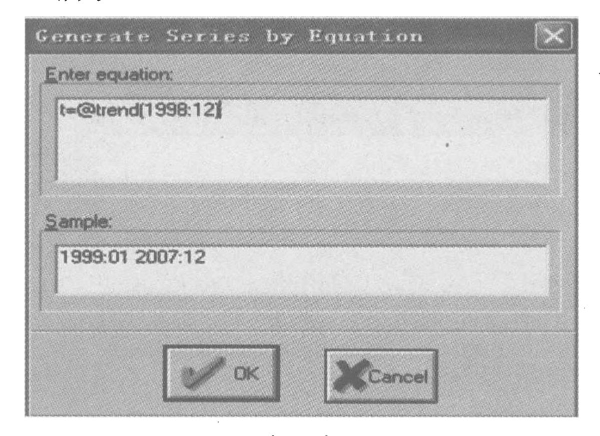


图 2—2　建立时间项 t 序列

3. 建立简单外推模型

（1）建立线性趋势模型

模型形式为：$y_t = c_1 + c_2 t$ 。从 Eviews 主选单中选“Quick/Estimate Equation”，

或者用新建对象的方法新建一个 Equation 对象，如图 2—3 所示。

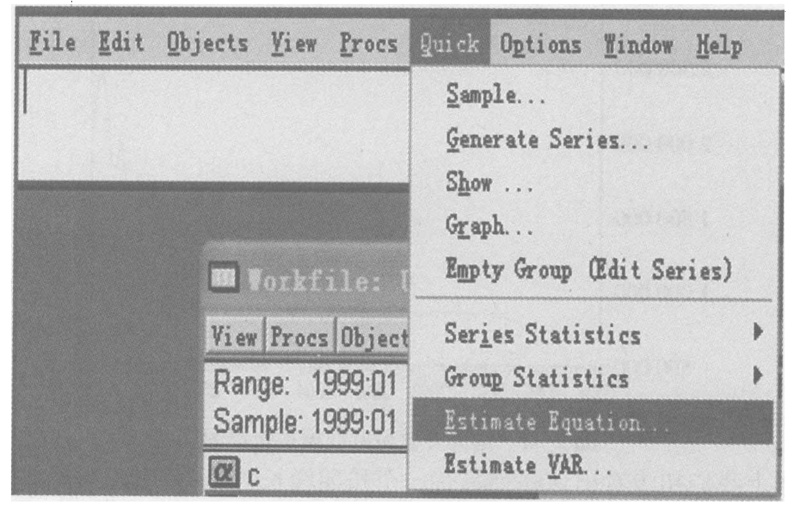


图 2—3　建立方程对象

在方程定义对话窗“Equantion Specification”选择框中输入“ck c t”，在“Estimation settings”选择框选择 LS 估计法和样本区间“1999:01 2007:12”，如图 2—4 所示。

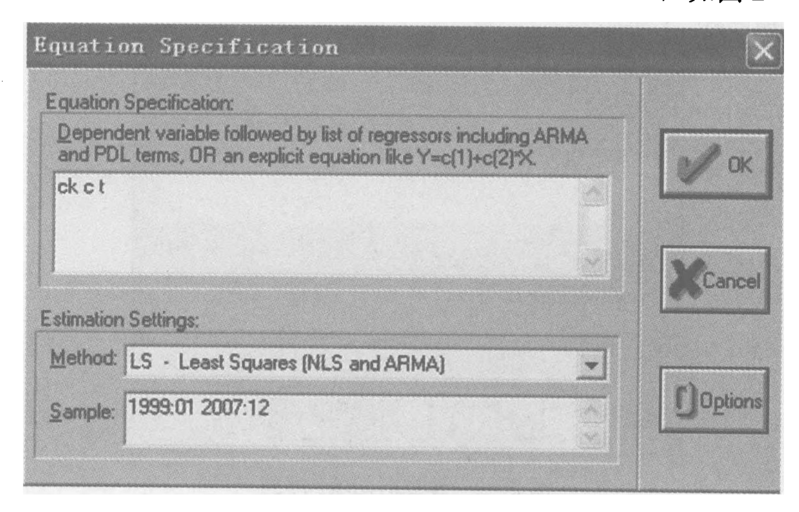


图 2—4　对方程进行定义

点击“OK”键后，得到如图 2—5 所示的估计结果，并将方程命名为 xxqs。

对应的回归表达式是

$$CK=1\,499\,969+9\,830.965\times T$$

$$(34.155)\quad(13.926)\qquad R^2=0.656,\ DW=0.144,\ F=202.5$$

其中：括号内的值是 t 统计量（下同）。

在方程窗口中选中“View/Actual,Fitted,Residual/ Actual,Fitted,Residual Graph”，得到真实值、拟合值与残差图，如图 2—6 所示。可以看到残差有明显的自相关，这与 DW=0.144 说明了同样的问题。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1499969.	43916.73	34.15485	0.0000
T	9830.965	705.9509	13.92585	0.0000

R-squared	0.648746	Mean dependent var	2030841.
Adjusted R-squared	0.645401	S.D. dependent var	378766.7
S.E. of regression	225548.8	Akaike info criterion	27.50898
Sum squared resid	5.34E+12	Schwarz criterion	27.55894
Log likelihood	-1469.730	F-statistic	193.9292
Durbin-Watson stat	0.131302	Prob(F-statistic)	0.000000

图 2—5　线性趋势模型方程输出结果

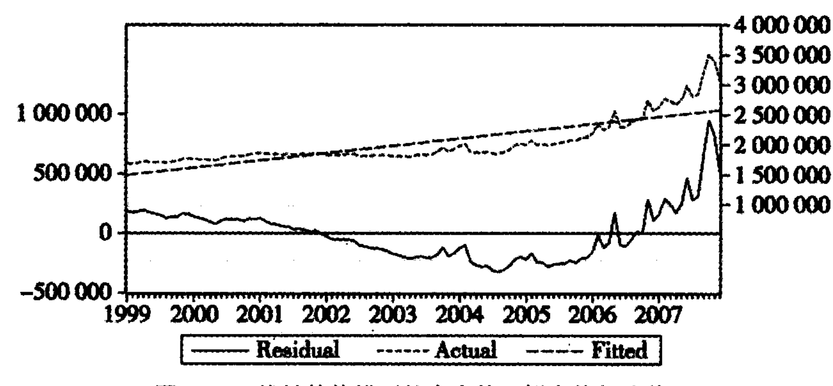


图 2—6　线性趋势模型的真实值、拟合值与残差图

利用线性趋势模型进行预测，方法是在方程窗口中选中 Forecast，出现如图 2—7 所示的对话框，在 Forecast name（预测序列命名）栏输入“xxqsf”用来保存线性趋势模型的预测值，S.E.（标准误）序列可以选填，Sample range for forecast（预测的样本范围）使用默认的样本范围“1999:01 2007:12”，Method（预测方法）选择“Static（静态预测）”，其他使用默认值。

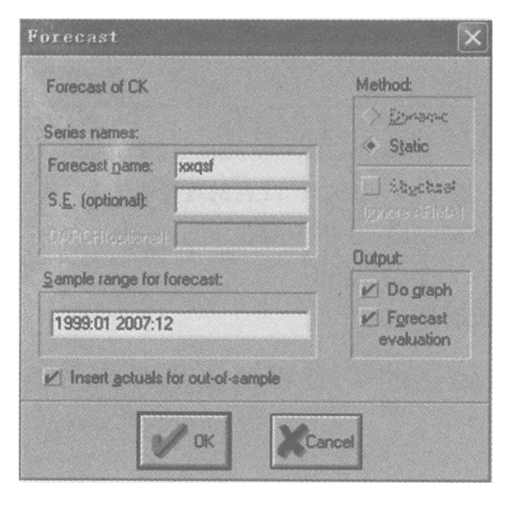


图 2—7　利用线性趋势模型进行预测

点击“OK”键，得到预测结果如图 2—8 所示。

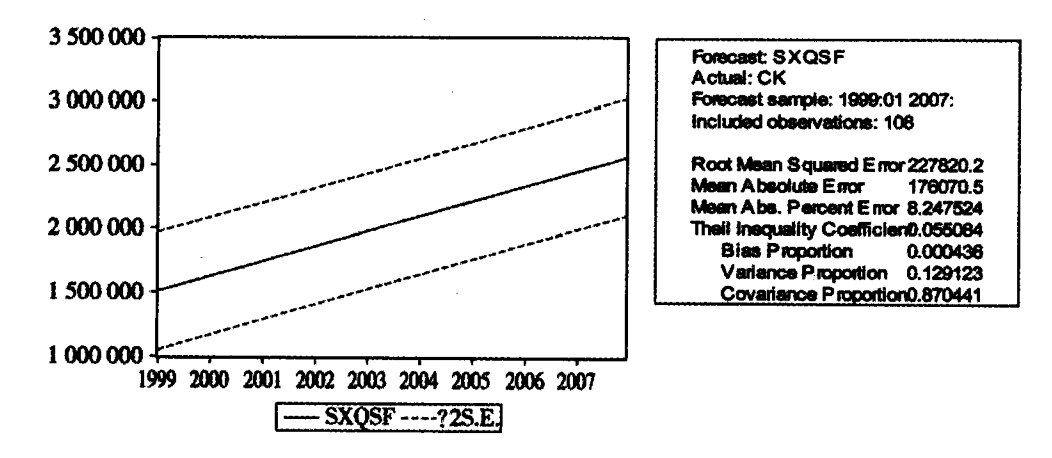


图 2—8　线性趋势模型预测图

图 2—8 中各种指标分别为：

Root Mean Squared Error（RMSE）：均方根误差。

Mean Absolute Error（MAE）：平均绝对误差。

Mean Abs.Percent Error（MAPE）：平均绝对百分比误差。

Theil Inequality Coefficient：希尔不等系数。

bias Proportion：偏差率。

Variance Proportion：方差率。

CovarianceProportion：协变率。

这里我们主要看 MAPE，它评价了整个样本期的整体预测效果。该模型（线性趋势模型）预测的平均绝对百分比误差为 8.24。

观察 ck 序列和 xxqsf 序列 2007 年 12 月的真实值与预测值，并比较，如图 2—9 所示。

obs	CK	SXQSF
2007:12	3075360.	2561713.

图 2—9　2007 年 12 月的真实值与预测值

预测误差为：$\eta = \dfrac{2\ 561\ 713 - 3\ 075\ 360}{3\ 075\ 360} = -0.167$。它评价了模型近期预测的效果。

（2）建立指数增长模型

模型形式为：$y_t = f(t) = Ae^{rt}$，方程两边取对数之后，得到回归方程为

$\log y_t = c_1 + c_2 t$

其中：$c_1=\log A$；$c_2=r$。建立方程对象，命名为 ZSZZ。在方程定义对话框输入“log（ck）c t”。得到的回归结果如图 2—10 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14.26565	0.017325	823.4341	0.0000
T	0.004548	0.000276	16.48181	0.0000

R-squared	0.719317	Mean dependent var	14.51351
Adjusted R-squared	0.716669	S.D. dependent var	0.167948
S.E. of regression	0.089397	Akaike info criterion	-1.973120
Sum squared resid	0.847128	Schwarz criterion	-1.923451
Log likelihood	108.5485	F-statistic	271.6501
Durbin-Watson stat	0.133367	Prob(F-statistic)	0.000000

图 2—10　指数增长模型方程输出结果

对应的回归表达式是

LOG（CK）=14.266+0.00455×T

（823.43）（16.48）　　　　R^2=0.719，DW=0.133，F=217.65

再做真实值、拟合值与残差图，如图 2—11 所示。可以看到残差有明显的自相关，这与 DW=0.133 说明了同样的问题。

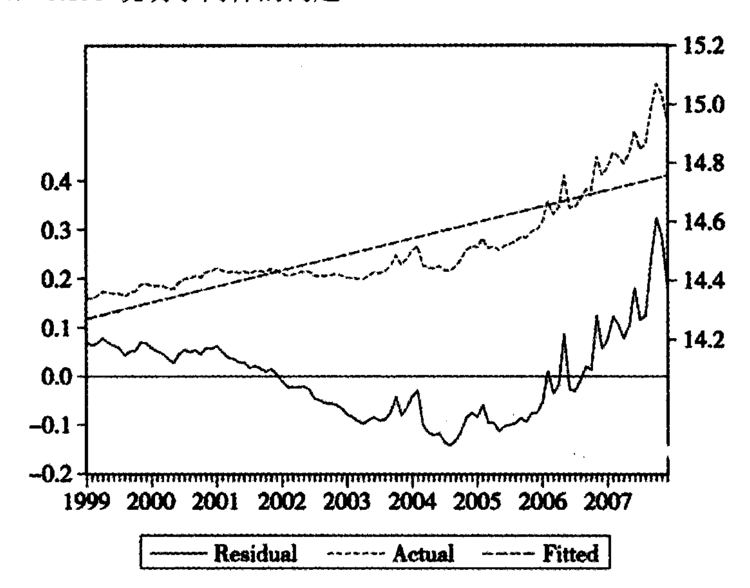


图 2—11　指数增长模型的真实值、拟合值与残差图

再利用指数增长模型进行预测，用 zszzf 来保存线性趋势模型的预测值，在 Method（预测方法）中选择“Static（静态预测）”，其他使用默认值，如图 2—12 所示。

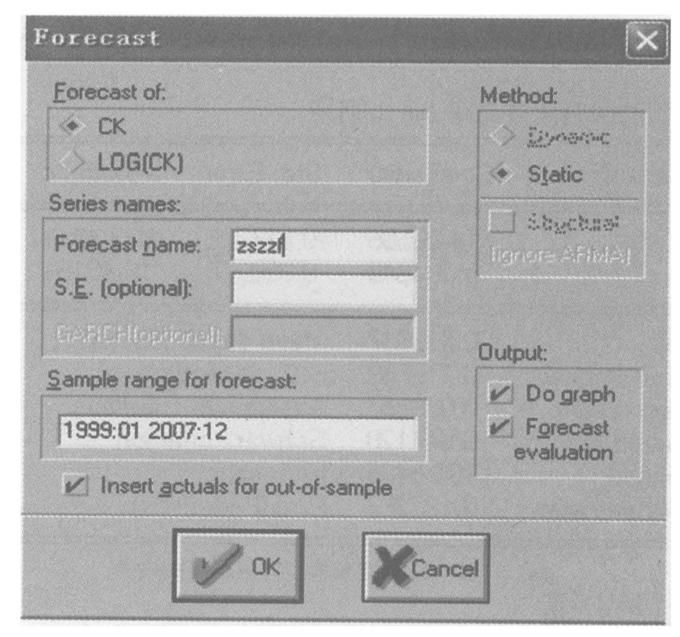


图 2—12 利用指数增长模型进行预测

这里多了一个预测对象，即可以对原序列 ck 做预测，也可以对 log（ck）序列做预测，这里我们选择对原序列 ck 做预测。预测结果如图 2—13 所示。该模型（指数增长模型）预测的平均绝对百分比误差（MAPE）为 7.17。

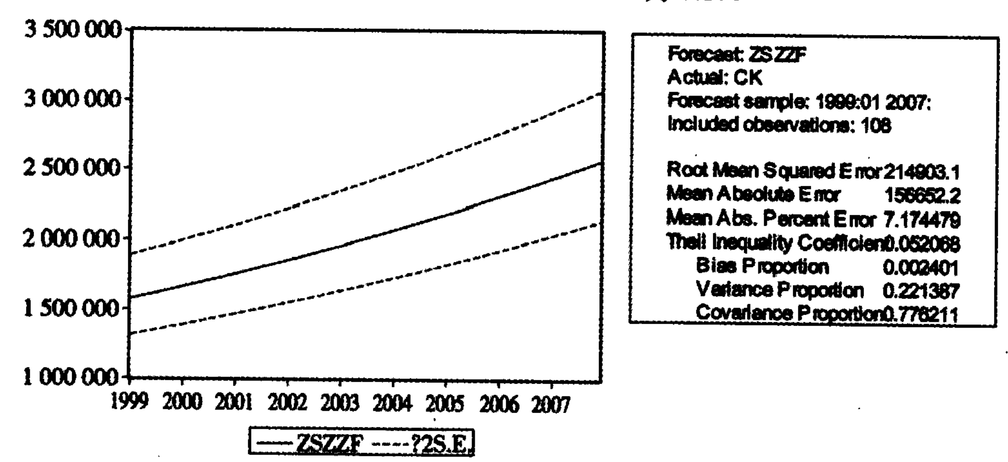


图 2—13 指数增长模型预测图

观察 ck 序列和 zszzf 序列 2007 年 12 月的真实值与预测值，并比较，如图 2—14 所示。

预测误差为：$\eta=\dfrac{2\ 563\ 319-3\ 075\ 360}{3\ 075\ 360}=-0.166$

obs	CK	ZSZZF
2007:12	3075360.	2563319.

图 2—14　2007 年 12 月的真实值与预测值

（3）建立自回归趋势模型

模型形式为：$y_t = c_1 + c_2 y_{t-1}$。建立方程对象，命名为 zhg。在方程定义对话框输入"ck c ck（-1）"。得到的回归结果如图 2—15 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2884.213	46381.80	0.062184	0.9505
CK(-1)	1.004959	0.022455	44.75417	0.0000
R-squared	0.950188	Mean dependent var		2043796.
Adjusted R-squared	0.949714	S.D. dependent var		390494.9
S.E. of regression	87566.81	Akaike info criterion		25.61671
Sum squared resid	8.05E+11	Schwarz criterion		25.66667
Log likelihood	-1368.494	F-statistic		2002.936
Durbin-Watson stat	2.004871	Prob(F-statistic)		0.000000

图 2—15　自回归趋势模型方程输出结果

常数项 t 统计量值为 0.062，相伴概率为 0.951，没有通过显著性检验，因此从方程中剔除，重新定义方程为 ck ck（-1）。得到回归结果如图 2—16 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CK(-1)	1.006332	0.004079	246.7042	0.0000
R-squared	0.950186	Mean dependent var		2043796.
Adjusted R-squared	0.950186	S.D. dependent var		390494.9
S.E. of regression	87154.39	Akaike info criterion		25.59805
Sum squared resid	8.05E+11	Schwarz criterion		25.62303
Log likelihood	-1368.496	Durbin-Watson stat		2.007382

图 2—16　调整后的自回归趋势模型方程输出结果

对应的回归表达式是

CK=1.0063×CK（-1）

（246.7）　　　　R^2=0.95

由于方程右边有被解释变量的滞后项 CK（-1），DW 检验失效。我们采用 LM 统计量进行检验（p=1），结果如图 2—17 所示。

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.678567	Probability	0.411948
Obs*R-squared	0.686921	Probability	0.407213

图 2—17 *LM*统计量检验结果

LM 统计量的相伴概率为 0.407，大于 0.05，可以认为在 5%的显著性水平上，残差没有一阶自相关。

再做真实值、拟合值与残差图，如图 2—18 所示。可以看到残差没有自相关，这与 *LM* 统计量检验说明了同样的问题，但有明显的异方差性。

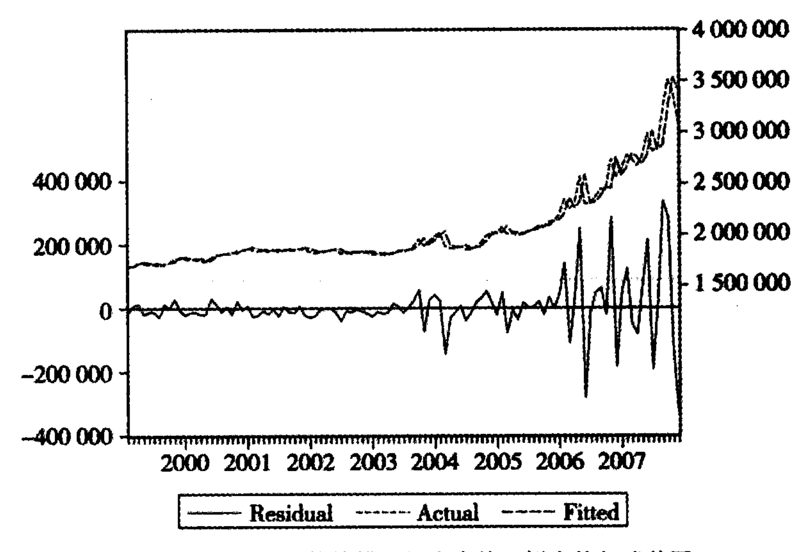


图 2—18 自回归趋势模型的真实值、拟合值与残差图

利用自回归趋势模型进行预测，用 zhgf 来保存线性趋势模型的预测值，Method（预测方法）选择 Static（静态预测），其他使用默认值，如图 2—19 所示。

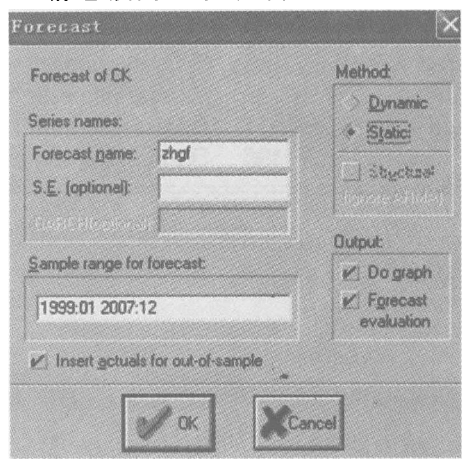


图 2—19 利用自回归趋势模型进行预测

预测结果如图 2—20 所示。该模型（自回归趋势模型）预测的平均绝对百分比误差（MAPE）为 1.98。

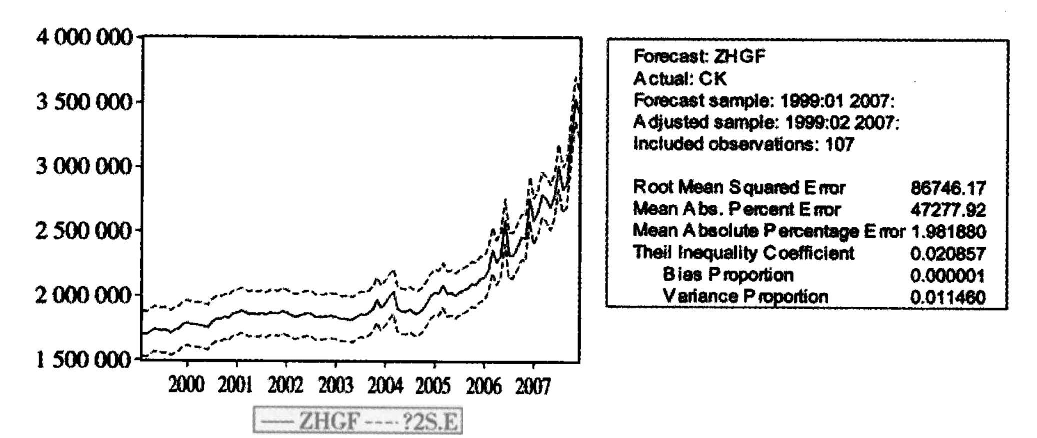


图 2—20　自回归趋势模型预测图

观察 ck 序列和 zhgf 序列 2007 年 12 月的真实值与预测值（如图 2—21 所示），并比较。

obs	CK	ZHGF
2007:12	3075360.	3411321.

图 2—21　2007 年 12 月的真实值与预测值

预测误差为：$\eta=\dfrac{3\ 411\ 321-3\ 075\ 360}{3\ 075\ 360}=0.1092$

（4）建立对数自回归趋势模型

模型形式为：$\log y_t = c_1 + c_2 \log y_{t-1}$。建立方程对象，命名为 dszhg。在方程定义对话框输入“log（ck）c log（ck（-1））”。得到回归结果如图 2—22 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.095072	0.282349	-0.336720	0.7370
LOG(CK(-1))	1.006938	0.019458	51.74851	0.0000

R-squared	0.962270	Mean dependent var	14.51513
Adjusted R-squared	0.961910	S.D. dependent var	0.167884
S.E. of regression	0.032765	Akaike info criterion	-3.980382
Sum squared resid	0.112724	Schwarz criterion	-3.930422
Log likelihood	214.9504	F-statistic	2677.908
Durbin-Watson stat	2.170305	Prob(F-statistic)	0.000000

图 2—22　对数自回归趋势模型方程输出结果

常数项 t 统计量值为-0.337，相伴概率为 0.737，没有通过显著性检验，因此从

方程中剔除，重新定义方程为 log（ck）log（ck（-1））。得到回归结果如图 2—23 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(CK(-1))	1.000387	0.000217	4602.046	0.0000

R-squared	0.962229	Mean dependent var	14.51513
Adjusted R-squared	0.962229	S.D. dependent var	0.167884
S.E. of regression	0.032628	Akaike info criterion	-3.997994
Sum squared resid	0.112846	Schwarz criterion	-3.973014
Log likelihood	214.8927	Durbin-Watson stat	2.154360

图 2—23　调整后的对数自回归趋势模型方程输出结果

对应的回归表达式是

LOG（CK）=1.000387×LOG（CK（-1））

（4602.046）　　　　　　R^2=0.96

由于方程右边有被解释变量的滞后项 LOG（CK（-1）），*DW* 检验失效。我们采用 *LM* 统计量进行检验（p=1），结果如图 2—24 所示。

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.833641	Probability	0.178605
Obs*R-squared	1.836482	Probability	0.175364

图 2—24　*LM* 统计量检验结果

LM 统计量的相伴概率为 0.175，大于 0.05，可以认为在 5%的显著性水平上，残差没有一阶自相关。

再做真实值、拟合值与残差图，如图 2—25 所示。可以看到残差没有自相关，这与 *LM* 统计量检验说明了同样的问题，但有明显的异方差性。

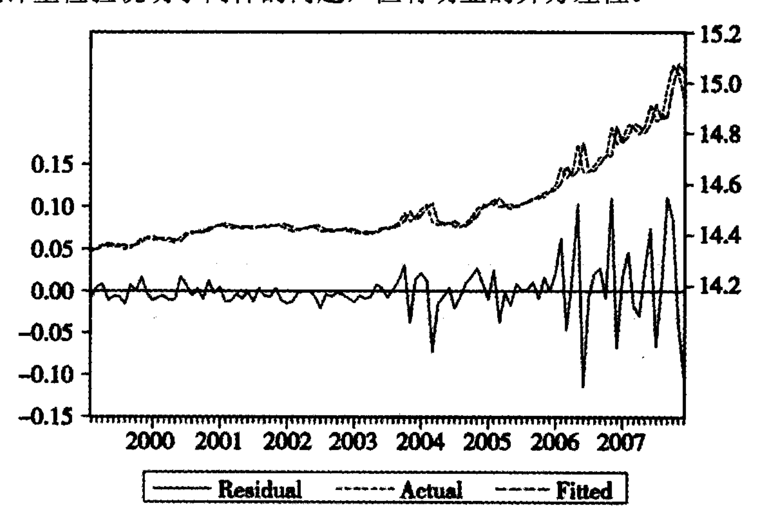


图 2—25　对数自回归趋势模型的真实值、拟合值与残差图

利用对数自回归趋势模型进行预测，用 dszhgf 来保存线性趋势模型的预测值，在 Method（预测方法）中选择“Static（静态预测）”，其他使用默认值，如图 2—26 所示。

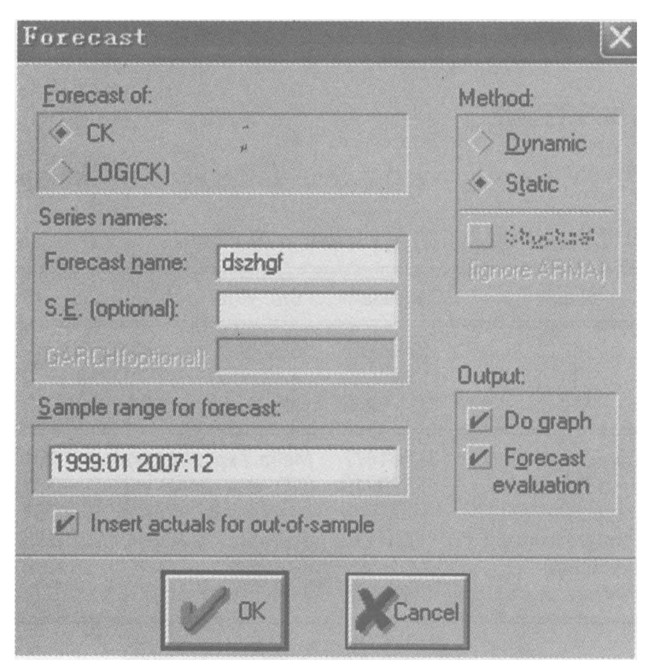


图 2—26　利用对数自回归趋势模型进行预测

预测结果如图 2—27 所示。该模型（对数自回归趋势模型）预测的平均绝对百分比误差（MAPE）为 1.969。

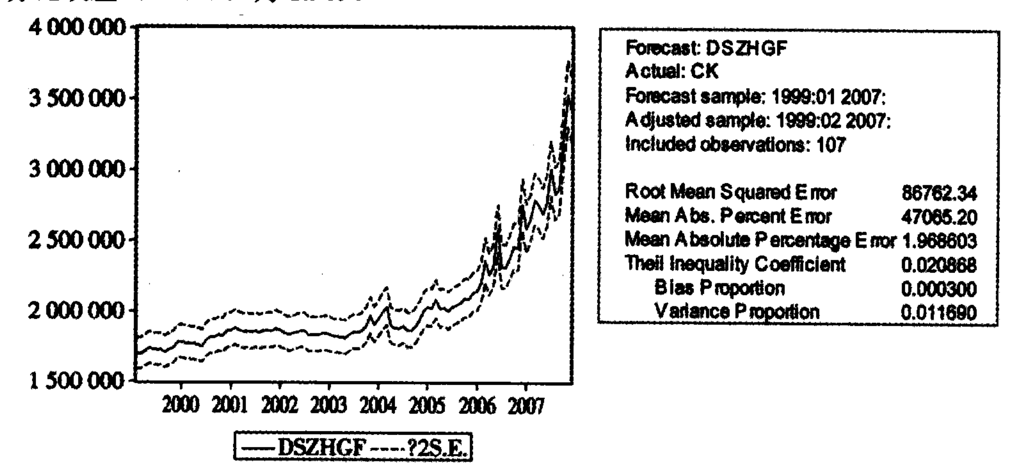


图 2—27　对数自回归趋势模型预测图

观察 ck 序列和 dszhgf 序列 2007 年 12 月的真实值与预测值，并比较，如图 2—28 所示。

obs	CK	DSZHGF
2007:12	3075360.	3409629.

图 2—28　2007 年 12 月的真实值与预测值

预测误差为：$\eta=\dfrac{3\ 409\ 629-3\ 075\ 360}{3\ 075\ 360}=0.1087$

（5）建立二次趋势模型

模型形式为：$y_t=c_1+c_2t+c_3t^2$。建立方程对象，命名为 ecqs。在方程定义对话框输入“ck c t t^2”。得到回归结果如图 2—29 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1913993.	39879.07	47.99494	0.0000
T	-13009.20	1688.889	-7.702813	0.0000
T^2	211.9446	15.01111	14.11918	0.0000

R-squared	0.881477	Mean dependent var	2040513.
Adjusted R-squared	0.879219	S.D. dependent var	390160.8
S.E. of regression	135594.8	Akaike info criterion	26.50011
Sum squared resid	1.93E+12	Schwarz criterion	26.57462
Log likelihood	-1428.006	F-statistic	390.4505
Durbin-Watson stat	0.410995	Prob(F-statistic)	0.000000

图 2—29　二次趋势模型方程输出结果

对应的回归表达式是

CK=1 913 993－13 009.20×T+211.9446×（T^2）

（74.99）　（-7.7）　　　（14.1）　　　　R^2=0.88，DW=0.41，F=390

再做真实值、拟合值与残差图，如图 2—30 所示。可以看到残差有明显的自相关，这与 DW=0.41 说明了同样的问题，而且存在异方差性。

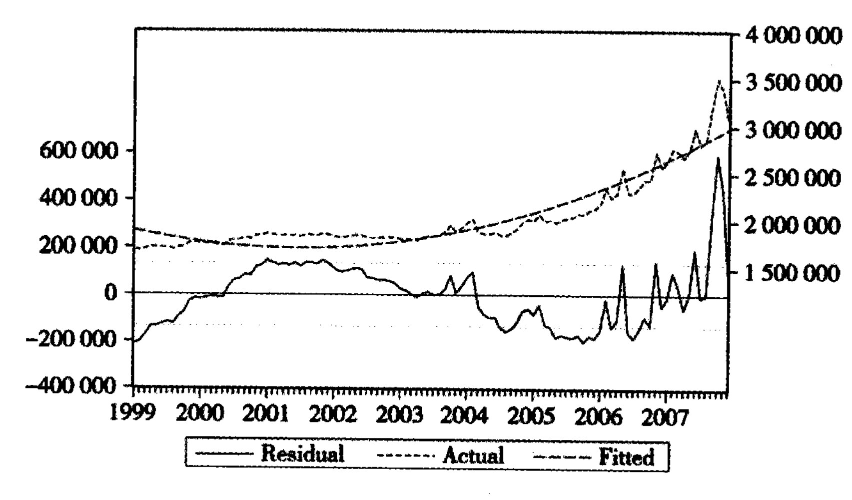


图 2—30　二次趋势模型的真实值、拟合值与残差图

利用二次趋势模型进行预测，用 ecqsf 来保存线性趋势模型的预测值，在 Method（预测方法）中选择“Static（静态预测）”，其他使用默认值，如图 2—31 所示。

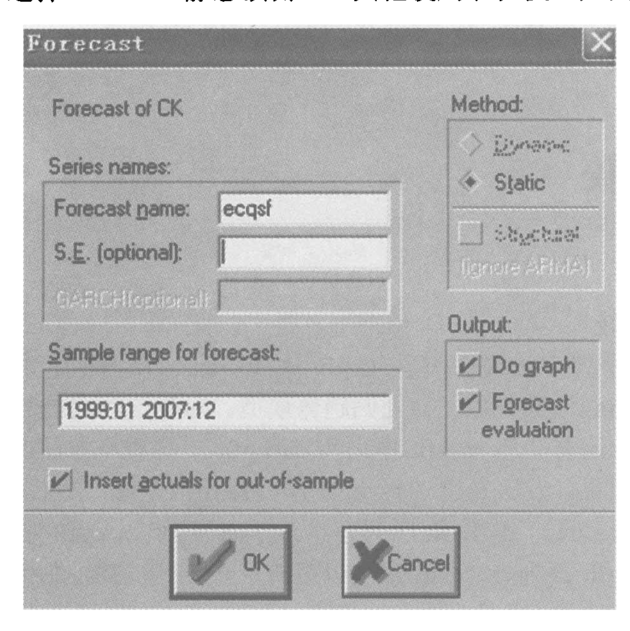


图 2—31　利用二次趋势模型进行预测

预测结果如图 2—32 所示。图中明显看出二次曲线的性质，该模型（对数自回归趋势模型）预测的平均绝对百分比误差（MAPE）为 5.04。

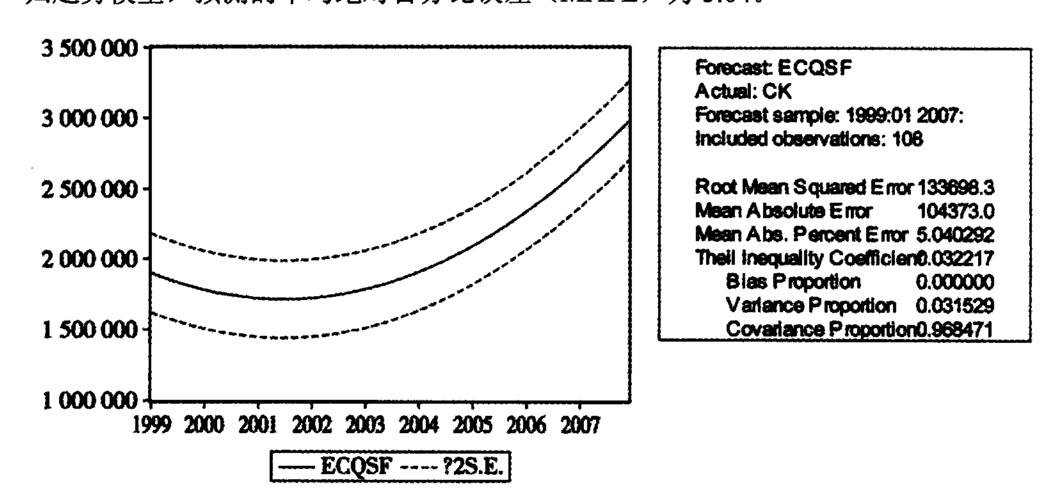


图 2—32　二次趋势模型的预测图

观察 ck 序列和 ecqsf 序列 2007 年 12 月的真实值与预测值，并比较，如图 2—33 所示。

预测误差为：$\eta=\dfrac{2\,981\,121-3\,075\,360}{3\,075\,360}=0.031$

obs	CK	ECQSF
2007:12	3075360.	2981121.

图 2—33 2007 年 12 月的真实值与预测值

4. 模型比较

（1）图形比较

建立 xxqsfit 序列、zszzfit 序列、zhgfit 序列、dszhgfit 序列、ecqsfit 序列，保存线性趋势模型、指数增长模型、自回归趋势模型、对数自回归趋势模型、二次趋势模型的拟合数据，需要注意 zszzfit 序列和 dszhgfit 序列要将方程拟合值取反对数的形式得到可比的结果。方法是建立 zszzfit1 序列保存拟合值，然后建立 zszzfit 序列，令 zszzfit=exp（zszzfit1）。dszhgfit 序列亦然。将 ck 序列、xxqsfit 序列、zszzfit 序列、zhgfit 序列、dszhgfit 序列、ecqsfit 序列以组的形式打开，并绘制折线图，得到如图 2—34 所示的结果。

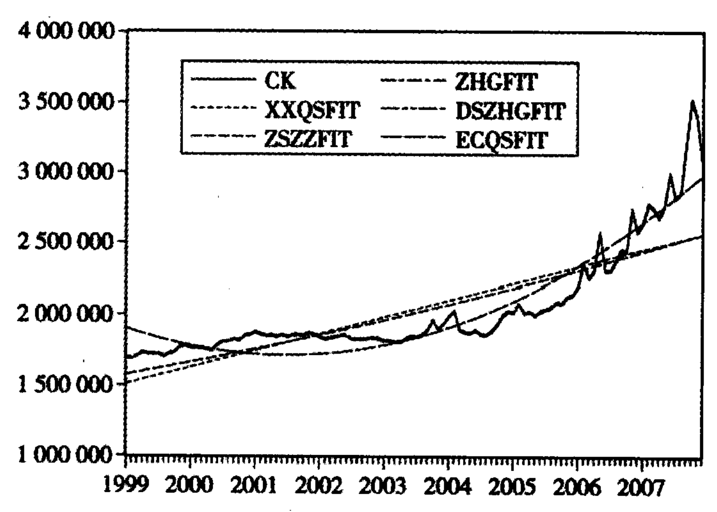


图 2—34 5 个模型拟合值与原序列折线图

可以看出自回归趋势模型、对数自回归趋势模型拟合效果最好。

将 ck 序列、xxqsf 序列、zszzf 序列、zhgf 序列、dszhgf 序列、ecqsf 序列以组的形式打开，并绘制折线图，得到如图 2—35 所示的结果。可以看出自回归趋势模型、对数自回归趋势模型预测效果最好。

（2）各种指标比较

将 5 个模型的 R^2，有无一阶自相关，MAPE，η 汇总到表 2—1。

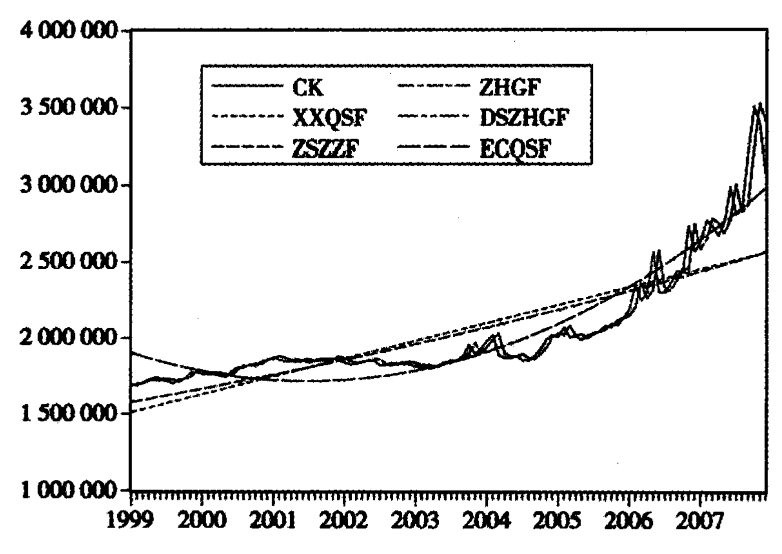


图 2—35　5 个模型预测值与原序列折线图

表 2—1　　5 个模型指标汇总表

模型	R^2	一阶自相关	MAPE	η
线性趋势模型	0.656	有	8.24	-0.167
指数增长模型	0.719	有	7.17	-0.166
自回归趋势模型	0.95	无	1.98	0.1092
对数自回归趋势模型	0.96	无	1.969	0.1087
二次趋势模型	0.88	有	5.04	0.031

从表 2—1 中我们可以看到，对数自回归趋势模型拟合和预测得最好，自回归趋势模型也很不错，其他 3 个模型拟合和预测效果不是很理想，这一点从上面的图形比较分析中也得到了相同的结论。但是，二次趋势模型的 2007 年 12 月预测误差仅为 0.031，是 5 个模型中最好的一个。这是为什么呢？我们从原序列图中看到从 2005 年左右开始，港元存款余额增长速度明显变快，而二次趋势模型正好拟合了这种趋势，因此更好的预测了 2007 年 12 月的数据。同时，我们怀疑是不是存在结构突变现象。由于本案例主要是为了介绍简单外推模型，结构突变的检验、拟合以及预测，留给读者自行完成。

5. 利用对数自回归趋势模型外推一期

从上面的分析，我们看出对数自回归趋势模型拟合和预测效果最理想，下面我

们利用对数自回归趋势模型对2008年1月港元存款余额进行外推预测。

（1）扩大工作文件范围

从Eviews主选单中选“Procs/Change workfile Range”，如图2—36所示。

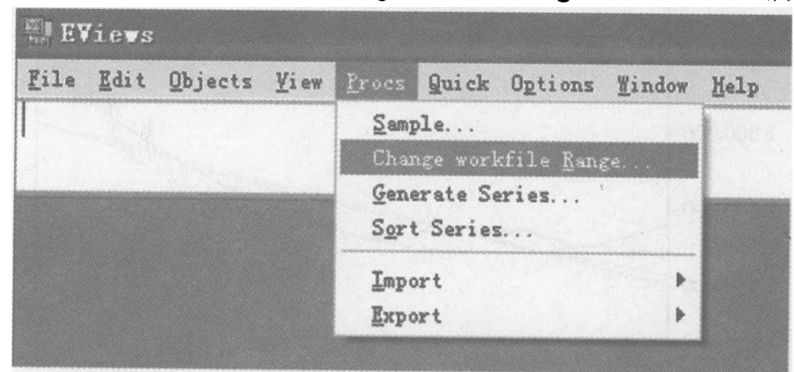


图2—36　选择修改工作文件范围

出现如图2—37所示的对话框，将End date修改为“2008:01”，点击“OK”。

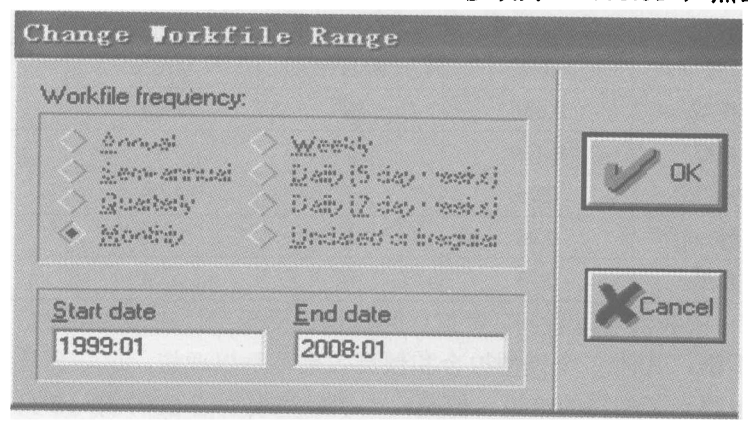


图2—37　将工作文件范围增加一期

（2）利用对数自回归趋势模型进行外推预测

打开dszhg方程对象（对数自回归趋势模型），在方程窗口中选中“Forecast”，出现如图2—38所示对话框，在Forecast name（预测序列命名）栏输入“dszhgf2”用来保存对数自回归趋势模型的预测值，Sample range for forecast（预测的样本范围）将默认的样本范围修改为“2008:01 2008:01”，即做点预测，Method（预测方法）选择“Static（静态预测）”，其他使用默认值。

（3）事后评价

2008年1月港元存款余额的真实值我们查到为3 020 639，与运用对数自回归趋势模型进行外推预测到的预测值3 093 179.8比较，预测误差为2.4%。应该说预测精度还是比较高的。

预测结果如图2—39所示，其中3093179.8为2008年1月港元存款余额的预测值，2 896 842.5～3 302 824.2为加减2倍标准误之后的范围。

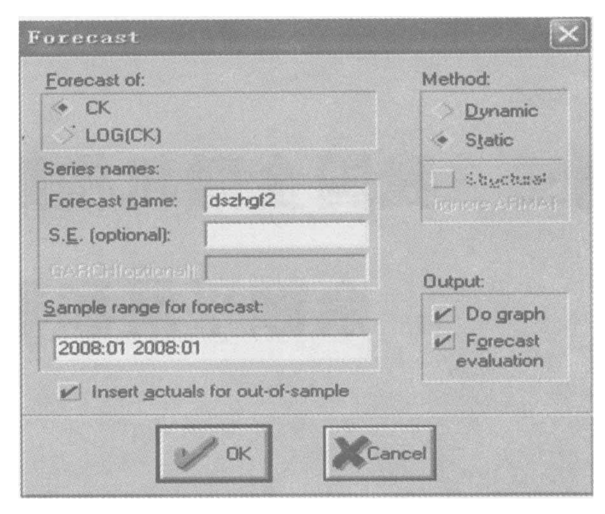


图 2—38　利用对数自回归趋势模型进行一期外推预测

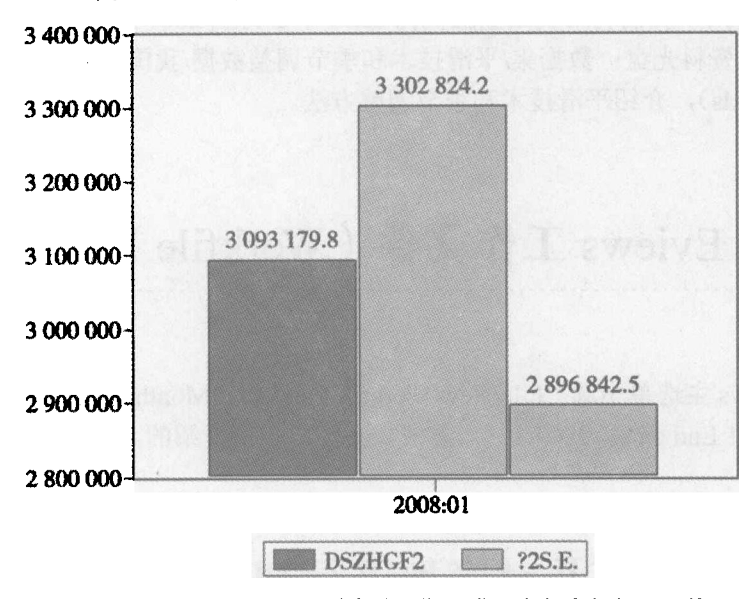


图 2—39　2008 年 1 月（样本外一期）港元存款余额的预测值

6. 小结

在时间、资源有限的情况下，当认为一个特定的时间序列有一个简单的趋势时，简单的外推模型作为快速、初步预测的方法是非常有用的，希望同学们能够熟练掌握和运用。

案例三　我国金融机构财政存款余额平滑和季节调整

在本案例中，我们利用我国金融机构2000年1月到2007年12月财政存款余额数据[①]（参见资料光盘：数据集/平滑技术和季节调整数据/我国金融机构财政存款余额月度数据.xls），介绍平滑技术和季节调整方法。

1. 创建Eviews工作文件（Workfile）

从Eviews主选单中选“File/New/Workfile”，选择“Monthly”选项，输入“Start date：2000:01 End date：2007:12”，方法如案例一中所介绍的。

2. 录入数据，并对序列进行初步分析

在Workfile窗口中选“Objects/New Object”，新建一个序列对象，命名为czck，用来保存财政存款余额数据，并将数据导入。该序列的折线图如图3—1所示。

从图3—1我们可以看到，财政存款余额序列有明显的上升趋势，并有明显的波动和季节因素，同时波动幅度随时间的增加而变大。

① 数据来源：中经网统计数据库。

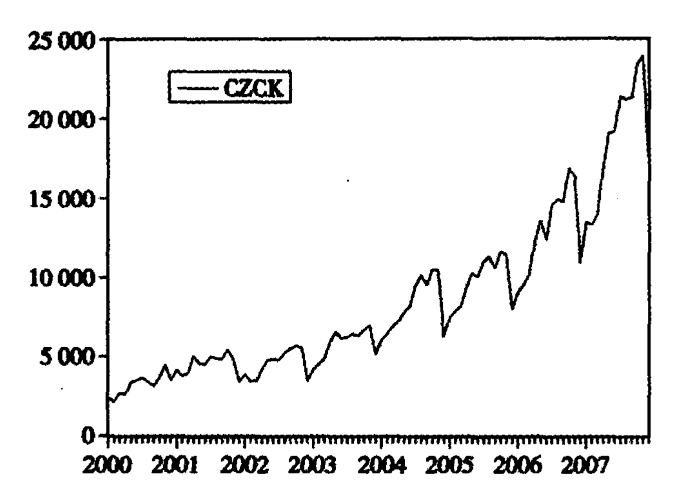


图 3—1　财政存款余额序列的折线图

从 Eviews 主选单中选“Quick/Generat Series”，输入“lnczck=log（czck）”，建立一个财政存款余额的对数序列（lnczck），该序列的折线图如图 3—2 所示。

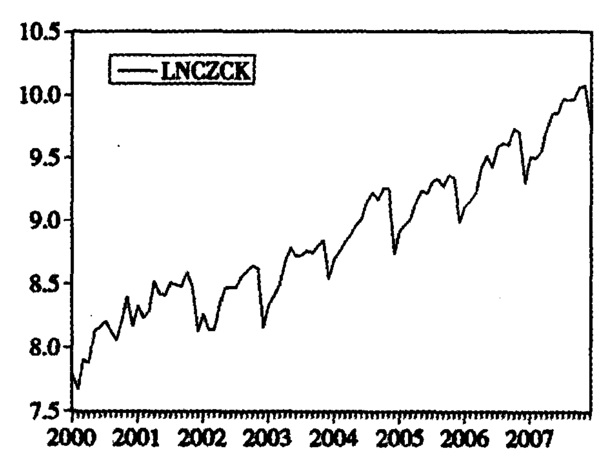


图 3—2　财政存款余额对数序列的折线图

从图 3—2 我们可以看到，财政存款余额对数序列具有明显的线性趋势，并有明显的波动和季节因素，但波动幅度随时间的变化改变不大，这主要归功于对数变换。

3. 简单移动平均平滑

这里我们对对数序列 lnczck 做 3 期和 7 期简单移动平均，并作比较。对原序列

czck 的简单移动平均留给读者自行完成。

（1）建立 3 期和 7 期简单移动平均序列

从 Eviews 主选单中选 “Quick/Generat Series”，输入 “lnczck3=（lnczck+lnczck（-1）+lnczck（-2））/3”，如图 3—3 所示。得到的 lnczck3 序列即 3 期简单移动平均序列。

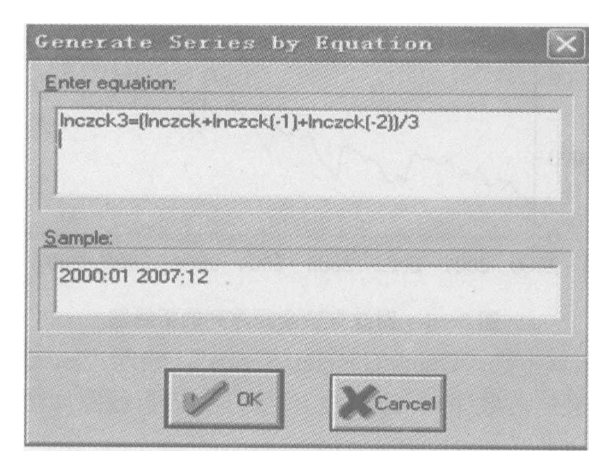


图 3—3　建立 3 期简单移动平均序列

同样建立 lnczck7 序列为 7 期简单移动平均序列，序列定义为 “lnczck7=（lnczck+lnczck（-1）+lnczck（-2）+lnczck（-3）+lnczck（-4）+lnczck（-5）+lnczck（-6））/7”，如图 3—4 所示。

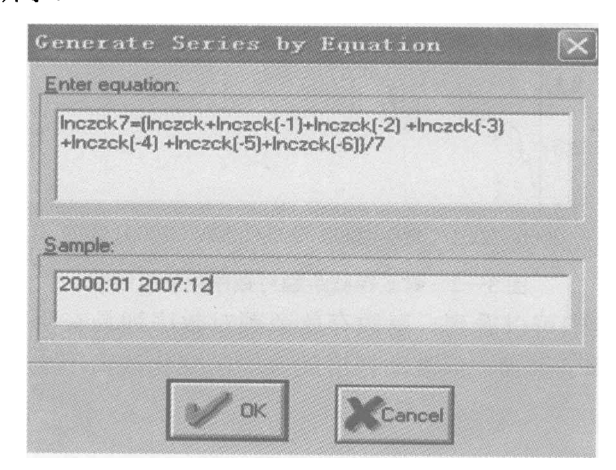


图 3—4　建立 7 期简单移动平均序列

（2）对数序列 lnczck 与 3 期、7 期简单移动平均序列比较

将对数序列 lnczck 与 lnczck3、lnczck7 序列以组的形式打开，并绘制折线图，得到如图 3—5 所示的结果。

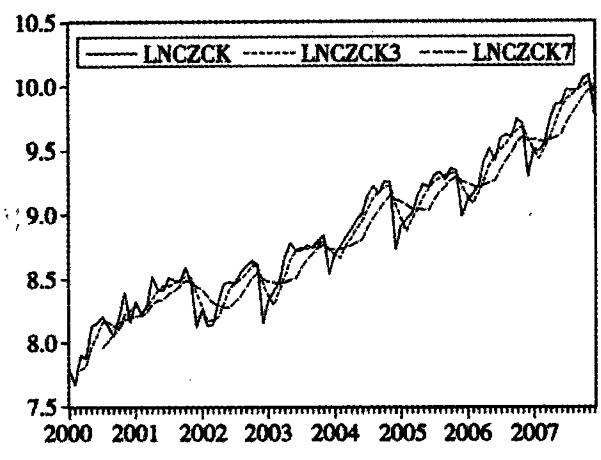


图 3—5　lnczck 与 lnczck3、lnczck7 序列比较

从图 3—5 我们可以看到，3 期、7 期简单移动平均都对原对数序列作了平滑，7 期简单移动平均的平滑程度更高，证明了简单移动平均的期数越多越平滑。同时，简单移动平均消除了一定的波动，甚至消除了一些季节变动影响，这使我们对于序列的长期趋势判断更加准确。但是，简单移动平均明显有滞后的现象。此处可以联系证券投资分析中的技术分析理论①。

4. 季节调整

通过对财政存款余额对数序列图（如图 3—2 所示）的分析，我们可以看到明显的季节变化。由于使用了对数变换，*L*、*S*、*C* 的关系接近于相互独立，因此我们选择加法模型进行季节调整。

打开 lnczck 序列，在序列窗口下选中“Procs/Seasonal Adjustment”，如图 3—6 所示。

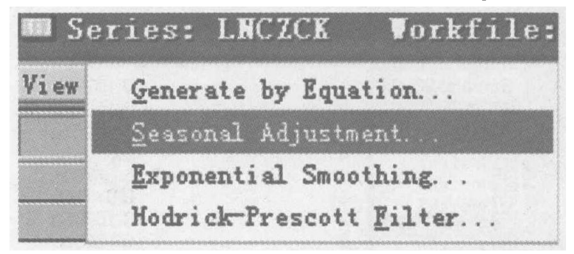


图 3—6　对财政存款余额对数序列进行季节调整

点击后出现如图 3—7 所示的季节调整定义对话框。

① 邢天才、王玉霞：《证券投资学》，大连，东北财经大学出版社，2003。

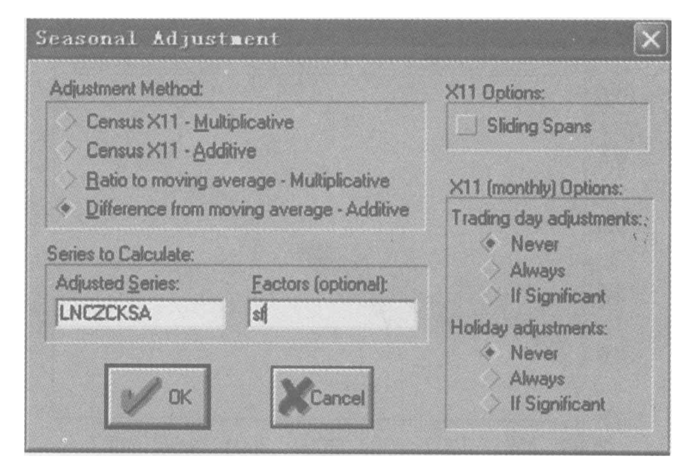


图 3—7 季节调整对话框

Eviews3.1 给出的季节调整方法(Adjustment Method)有四种，即 X11 乘法(Census X11-Multiplicative)、X11 加法(Census X11-Additive)、移动平均乘法(Ratio to moving average-Multiplicative）和移动平均加法（Difference from moving average- Additive)，这里我们选择移动平均加法。调整后序列（Adjusted Series）的名称，系统默认为在被季节调整序列名后面加 SA，我们不作改动。季节因子（Factors）可以选择性保存，我们选择将季节因子保存在序列 sf 中，即在该栏中输入 sf。由于没有使用 X11 方法，右侧的 X11 选项可以忽略。

点击“OK”后，得到如图 3—8 所示的结果，可以看到一月份的季节因子是-0.132653，二月份的季节因子是-0.152978 等。同时工作文件窗口出现季节调整后序列 lnczcksa 和季节因子序列 sf。

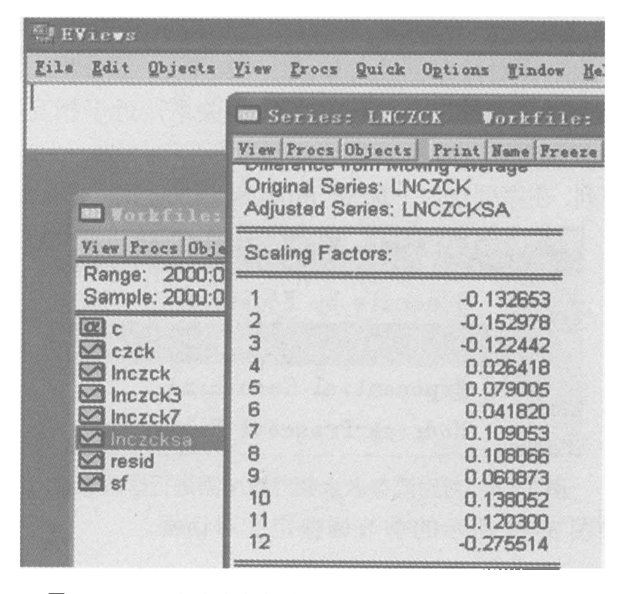


图 3—8 对财政存款余额对数序列季节调整结果

季节因子 sf 序列如图 3—9 所示，它是一个以 12 为周期的循环序列，相同月份的值相同。

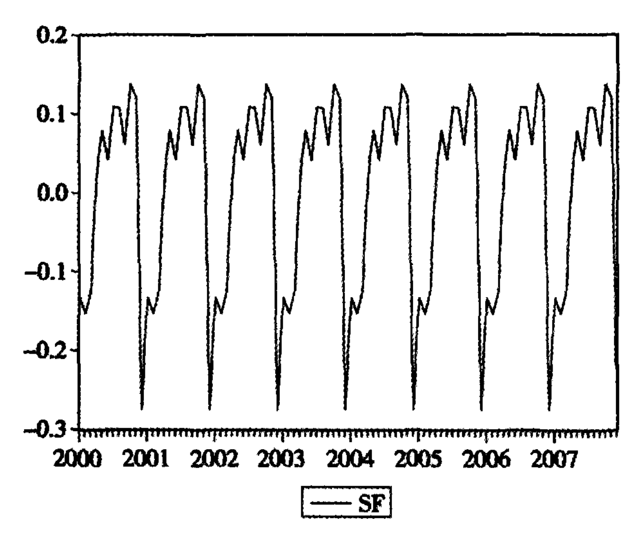
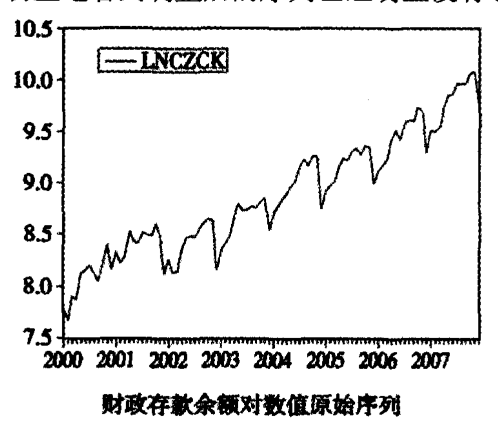
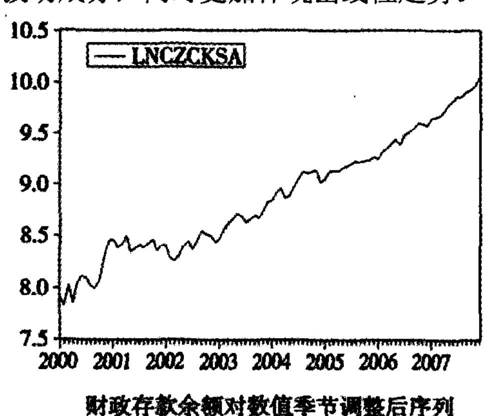


图 3—9　季节因子序列图

比较财政存款余额对数序列季节调整前后的序列图，如图 3—10 所示。可以很明显地看到调整后的序列已经明显没有季节波动成分，同时更加体现出线性趋势。



财政存款余额对数值原始序列



财政存款余额对数值季节调整后序列

图 3—10　财政存款余额对数序列季节调整前后比较图

5. 指数平滑

（1）平滑财政存款余额对数序列

打开财政存款余额对数序列 lnczck，在序列窗口下选中“Procs/Exponential

smoothing”，如图 3—11 所示。

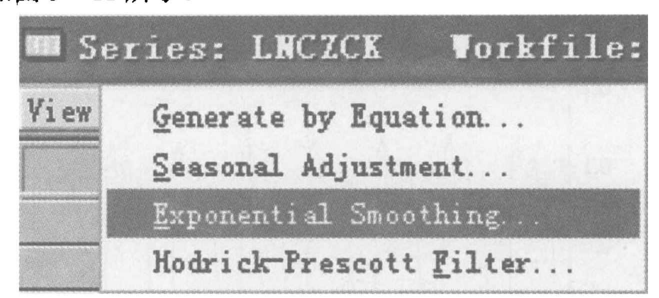


图 3—11 对财政存款余额对数序列进行指数平滑

点击后出现如图 3—12 所示的对话框。

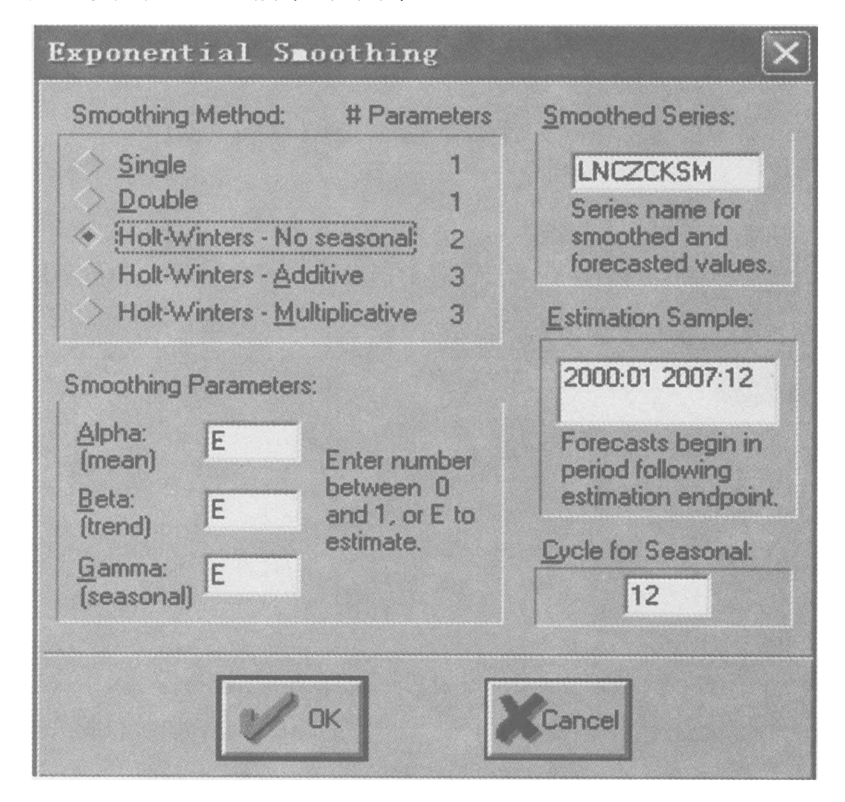


图 3—12 指数平滑对话框

Eviews3.1 给出的指数平滑方法（smoothing method）有五种，即一次指数平滑（Single Exponential Smoothing）、二次指数平滑（Double Exponential Smoothing）、Holt-Winters 非季节性模型（Holt-Winters-No Seasonal）、Holt-Winters 加法模型（Holt-Winters-Additive）和 Holt-Winters 乘法模型（Holt-Winters-Multiplicative）。

平滑系数（smoothing parameters）包括 Alpha（α），Beta（β），Gamma（γ），具体含义可以参看实验三的实验准备知识部分，参数值可以由用户自己指定也可由系统自动确定。如果用户指定，可以将 E 修改为需要的值，当然必须是 0 到 1 之间

的数。

平滑后序列名称（smoothed series name）系统默认为在被平滑序列名后加 SM，也可由用户自行命名。

样本范围（estimation sample）系统默认为工作文件的样本范围，同样也可以由用户自行修改。

季节循环（cycle for seasonal），是 Holt-Winters 加法模型和 Holt-Winters 乘法模型中使用的，季度数据填 4，月度数据填 12。

通过对财政存款余额对数序列图（如图 3—2 所示）的分析，我们可以看到明显的加法型季节波动和线性趋势，因此我们选择 Holt-Winters 加法模型，其他选择默认值。

点击“OK”后，得到如图 3—13 所示的结果。同时，工作文件中新增加 lnczcksm 序列。

Series: LNCZCK Workfile: 财政存款平滑和季

View | Procs | Objects | Print | Name | Freeze | Sample | Genr | Sheet

Method: Holt-Winters Additive Seasonal
Original Series: LNCZCK
Forecast Series: LNCZCKSM

Parameters:	Alpha		0.8500
	Beta		0.0000
	Gamma		0.0000
Sum of Squared Residuals			0.396289
Root Mean Squared Error			0.064250
End of Period Levels:	Mean		10.03036
	Trend		0.020987
	Seasonals:	2007:01	-0.137123
		2007:02	-0.171783
		2007:03	-0.117698
		2007:04	0.007960
		2007:05	0.082530
		2007:06	0.050225
		2007:07	0.111704
		2007:08	0.107358
		2007:09	0.063798
		2007:10	0.140309
		2007:11	0.125152
		2007:12	-0.262431

图 3—13 对财政存款余额对数序列指数平滑结果

可以看到系统选取的平滑系数值为 $\alpha=0.85$，$\beta=0$，$\gamma=0$。残差平方和（sum of squared residuals）为0.396289，均方根误差（root mean squared error）为0.064250，样本最后期的截距（mean）为10.03036，斜率（trend）为0.020987。季节因子（seasonals），列出了样本最后一年的加法季节因子。

可知，预测公式为：

$$\tilde{y}_{T+k}=10.03036+0.020987k+S_{T+k-s}\quad 对于所有k\geq 1 \tag{3.1}$$

（2）利用指数平滑方法预测财政存款余额

指数平滑方法的一个重要用途就是进行短期预测。用指数平滑方法进行预测也十分方便，只要在平滑之前将样本的范围扩大到预测期即可。以上面的财政存款余额序列为例，将样本范围扩大到2008年3月，那么2008年1月、2月、3月的值就是预测值。

首先，修改工作文件的范围，末期到2008年3月，方法在案例二中已经介绍，这里不再赘述。然后，将指数平滑对话框中的样本期修改为“2000:01 2008:03”。为了与前面平滑后序列lnczcksm区别，将平滑后序列名称改为lnczckf，表示预测值（如图3—14所示）。

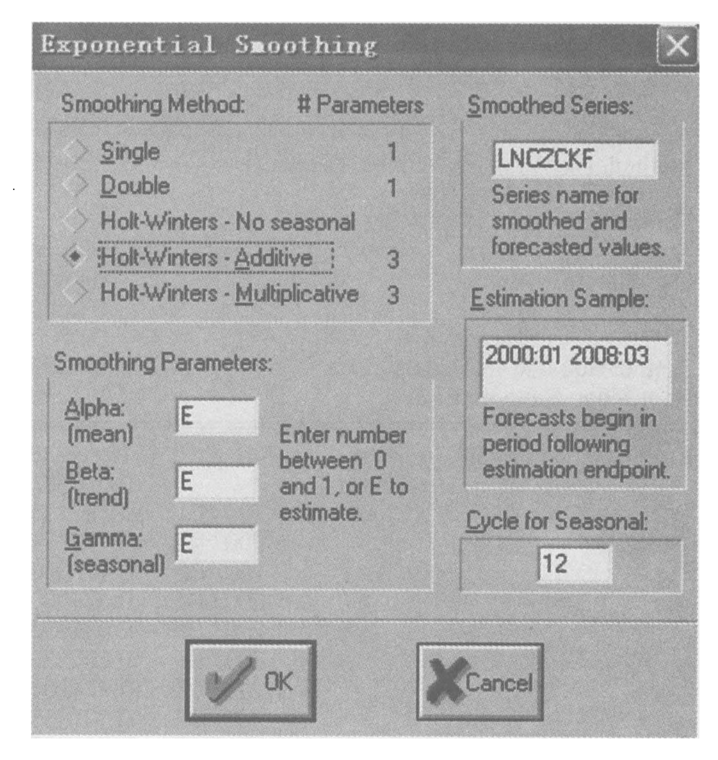


图3—14 指数平滑预测对话框

图3—15是财政存款余额对数序列与其预测值的比较图。

最后，我们建立序列czckf，令czckf=exp（lnczckf），得到财政存款余额原序列的预测值。

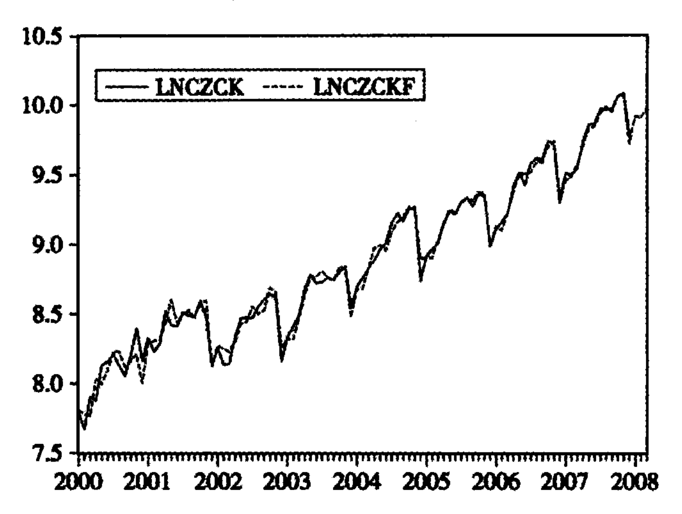


图 3—15　财政存款余额对数序列与其预测值比较图

图 3—16 列出了财政存款余额预测值。

obs	CZCK	CZCKF
2007:11	23923.30	23557.46
2007:12	17632.46	16542.70
2008:01	NA	20215.82
2008:02	NA	19941.30
2008:03	NA	21495.95

图 3—16　财政存款余额预测值

图 3—17 是财政存款余额原序列与其预测值的比较图。

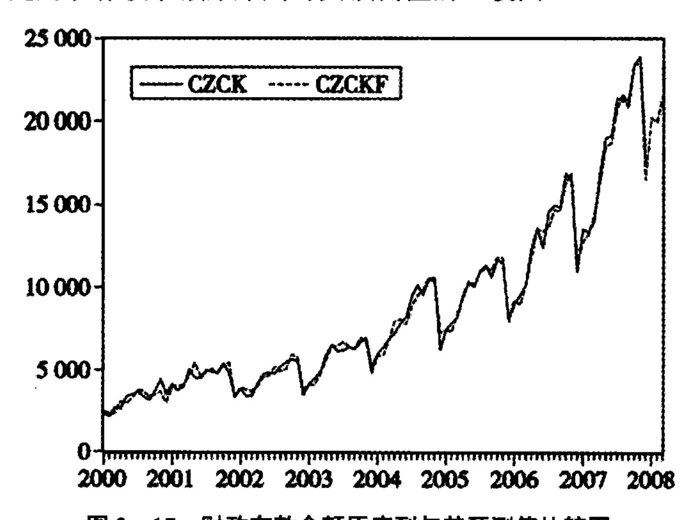


图 3—17　财政存款余额原序列与其预测值比较图

从图 3—15 和图 3—17，我们可以看到，利用指数平滑方法进行的预测很好的保留和继续了被预测序列的趋势和波动规律。从短期预测的角度看，指数平滑方法是一种非常有效和方便的预测方法。

案例四　我国金融机构储蓄存款余额序列的特征观察

在本案例中，我们利用我国金融机构1992年1月到2007年12月储蓄存款余额数据[①]（参见资料光盘：数据集/随机时间序列的特征观察数据/我国金融机构储蓄存款余额月度数据.xls），介绍随机时间序列特征的观察方法。

1. 创建Eviews工作文件（Workfile）

从Eviews主选单中选File/New/Workfile，选择Monthly选项，输入“Start date：1992:01 End date：2007:12”，方法如案例一中所介绍的。

2. 录入数据，并对序列进行初步分析

在Workfile窗口中选“Objects/New Object”，新建一个序列对象，命名为cxck，用来保存储蓄存款余额数据，并将数据导入。该序列的折线图如图4—1所示。

从图4—1我们可以看到，储蓄存款余额序列有明显的上升趋势，但没有明显的波动和季节因素。

[①] 数据来源：中经网统计数据库。

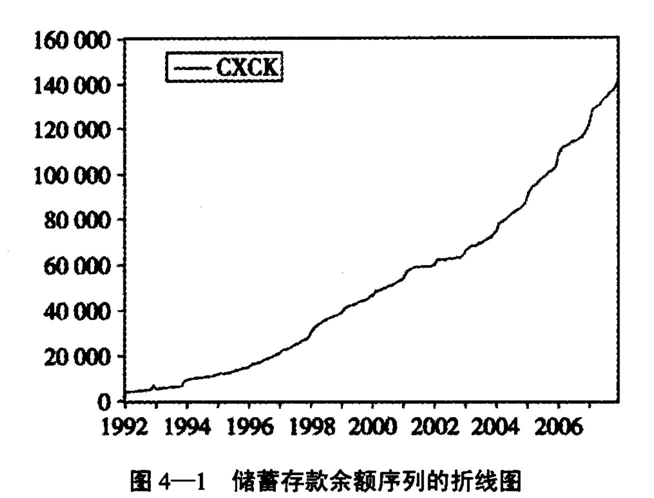


图 4—1　储蓄存款余额序列的折线图

3. 序列特征的观察

（1）原序列的相关分析图观察

打开 cxck 序列，在序列窗口下选中“View/Correlogram”，如图 4—2 所示。

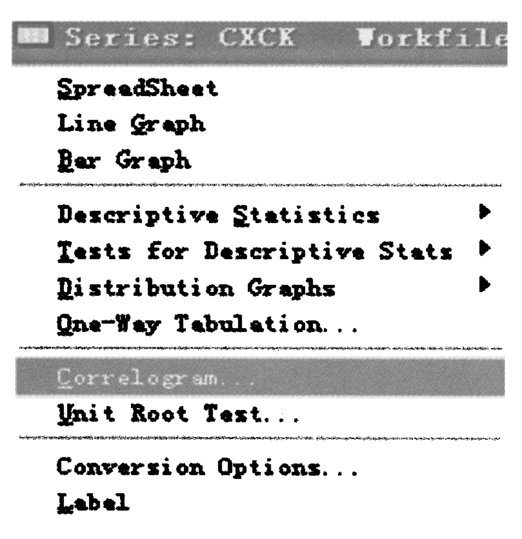


图 4—2　观察储蓄存款余额序列相关分析图

点击后出现如图 4—3 所示的对话框。其中，默认 level 表示不做差分，1st difference 表示做一次差分，2nd difference 表示做二次差分。这里我们观察原序列的相关分析图，选择 level。Lags to include 为定义计算自相关系数的最大滞后阶数 K，

这里我们取 15。

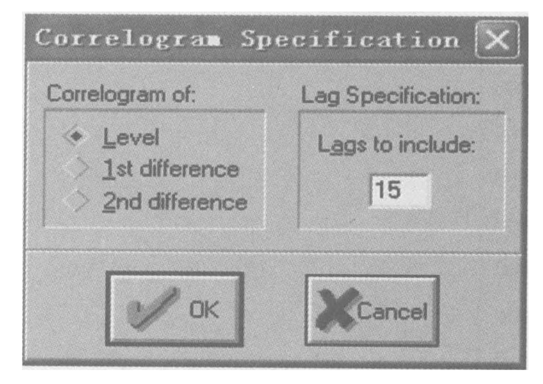


图 4—3　相关分析图对话框

点击“OK”后，得到如图 4—4 所示的结果。该图左半部分为序列的自相关和偏自相关系数图，虚线为 5%显著性水平下的置信带，右半部分的五列数值分别为：滞后期 K、自相关系数 r_k、偏自相关系数 $\varphi_{k,k}$、Q 统计量和相伴概率。可以看出，储蓄存款余额原序列的自相关系数随着 K 增加缓慢下降，但一直到 15 期也没有落入置信区间，说明该序列非平稳。同时很难看出季节性成分。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.981	0.981	187.86	0.000
		2	0.963	0.000	369.79	0.000
		3	0.945	-0.007	545.87	0.000
		4	0.927	-0.015	716.10	0.000
		5	0.909	-0.005	880.61	0.000
		6	0.891	-0.013	1039.4	0.000
		7	0.872	-0.008	1192.6	0.000
		8	0.854	-0.005	1340.4	0.000
		9	0.836	-0.013	1482.7	0.000
		10	0.818	-0.015	1619.6	0.000
		11	0.800	-0.009	1751.1	0.000
		12	0.783	0.034	1877.9	0.000
		13	0.766	-0.008	2000.1	0.000
		14	0.750	-0.006	2117.7	0.000
		15	0.733	-0.005	2230.8	0.000

图 4—4　储蓄存款余额原序列相关分析图

（2）差分序列的相关分析图观察

为了确定储蓄存款余额序列是否为一个齐次非平稳序列（如果是，需要差分多少次才能平稳），我们观察差分序列的相关分析图。

观察一次差分序列相关分析图。如图 4—3 所示，我们选择“1st difference”。同时为了观察是否有季节特征，自相关系数的最大滞后阶数 K，我们取 36。得到的结果如图 4—5 所示。

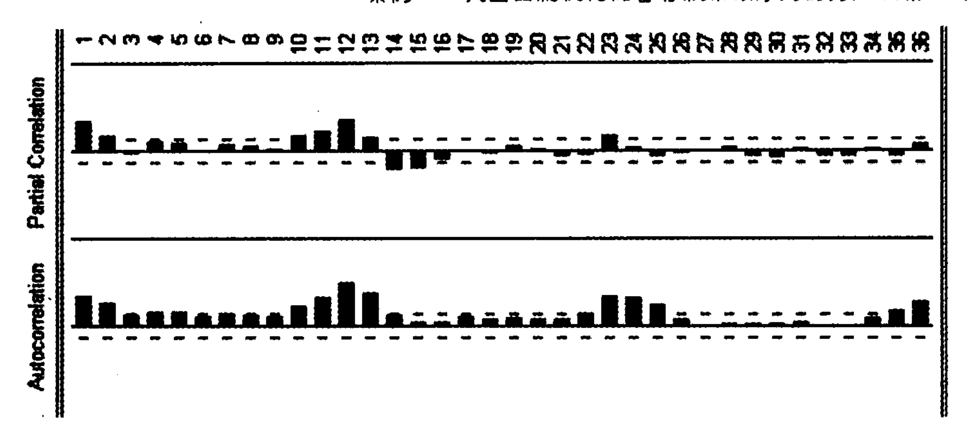


图 4—5　储蓄存款余额一次差分序列相关分析图

可以看出一次差分后的序列虽然近似平稳，但样本自相关系数 r_k，在 k=12，24，36 有明显的峰值，这充分反映出季节性。从图 4—1 和图 4—4 中，我们很难发现序列的季节特征，这主要是由于原序列的强趋势性掩盖了季节性。可见，相关分析图是我们分析序列特征的一个非常有用的工具。

（3）剔除季节成分后的相关分析图观察

由于季节成分的存在，可能会使我们对储蓄存款余额序列的分析存在偏差，因此我们利用季节调整手段剔除季节成分后，再观察序列。这里我们用移动平均加法模型对序列进行季节调整，方法同案例三，得到的序列命名为 cxcksa。

观察储蓄存款余额季节调整后序列（cxcksa）的相关分析图。如图 4—3 中，我们选择 level。最大滞后阶数 K，我们取 15。得到的结果如图 4—6 所示。可以看出，储蓄存款余额季节调整后序列的自相关系数随着 K 增加缓慢下降，但一直到 15 期也没有落入置信区间，说明还是非平稳序列。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.981	0.981	187.62	0.000
		2	0.963	-0.007	369.57	0.000
		3	0.944	-0.005	545.38	0.000
		4	0.926	-0.012	715.28	0.000
		5	0.908	-0.003	879.44	0.000
		6	0.890	-0.011	1037.9	0.000
		7	0.871	-0.005	1190.8	0.000
		8	0.854	-0.004	1338.3	0.000
		9	0.836	-0.010	1480.4	0.000
		10	0.818	-0.011	1617.2	0.000
		11	0.800	-0.007	1748.8	0.000
		12	0.783	0.025	1876.6	0.000
		13	0.766	-0.014	1997.8	0.000
		14	0.749	-0.011	2115.3	0.000
		15	0.733	-0.003	2228.3	0.000

图 4—6　储蓄存款余额季节调整后序列相关分析图

（4）季节调整后差分序列的相关分析图观察

观察季节调整后序列（cxcksa）的一次差分序列相关分析图。如图 4—3 中，选择“1st difference”。最大滞后阶数 K，我们还取 15。得到的结果如图 4—7 所示。可以看出 cxcksa 的一次差分序列还是不平稳。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.295	0.295	16.911	0.000
		2	0.353	0.292	41.283	0.000
		3	0.265	0.126	55.073	0.000
		4	0.316	0.170	74.772	0.000
		5	0.293	0.129	91.824	0.000
		6	0.236	0.029	102.91	0.000
		7	0.261	0.076	116.58	0.000
		8	0.259	0.080	130.06	0.000
		9	0.241	0.039	141.82	0.000
		10	0.282	0.104	158.04	0.000
		11	0.300	0.120	176.42	0.000
		12	0.393	0.212	208.20	0.000
		13	0.338	0.120	231.79	0.000
		14	0.156	-0.168	236.83	0.000
		15	0.143	-0.157	241.13	0.000

图 4—7　cxcksa 的一次差分序列相关分析图

我们再来观察季节调整后序列（cxcksa）的二次差分序列相关分析图。如图 4—3 中，选择 2nd difference。这次为了观察是否还存在季节成分，自相关系数的最大滞后阶数 K，我们取 36。得到的结果如图 4—8 所示。

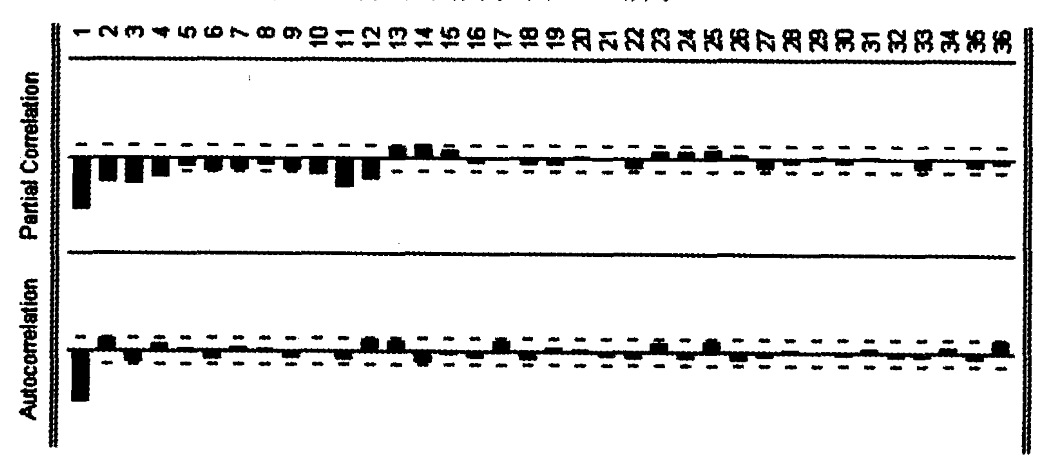


图 4—8　cxcksa 的二次差分序列相关分析图

从图 4—8 我们可以看出，cxcksa 的二次差分序列的自相关系数迅速下降为 0（从第二期开始），并且在 k=12，24，36 处，样本自相关系数 r_k 与 0 没有显著区别。可

以判断季节调整后的二次差分序列是一个没有季节成分的平稳序列。即 cxcksa 是一个二次齐次非平稳序列，平稳化需要的差分次数为 2 次。



图 4—9 是 cxcksa 的二次差分序列折线图，图形也支持上述结论。

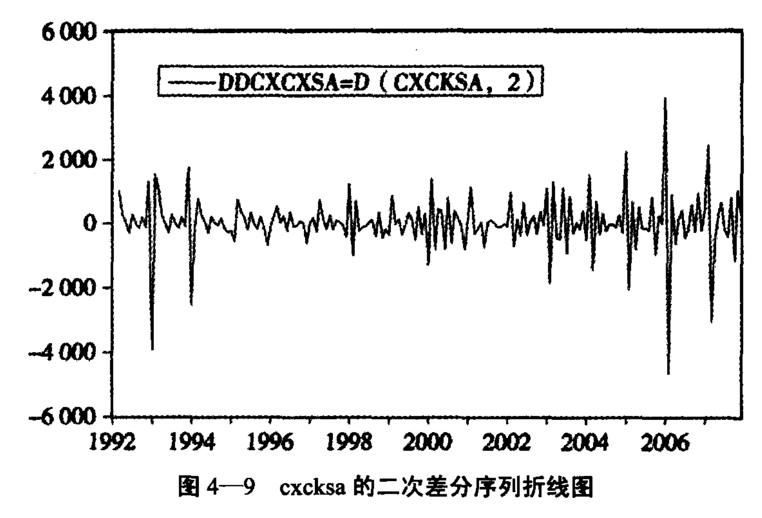


图 4—9 cxcksa 的二次差分序列折线图

4. 小结

我们在实际中遇到的时间序列很少是平稳序列。但是大部分经济、金融和商业数据是齐次非平稳序列。相关分析图可以确定一个序列是否平稳或确定一个齐次非平稳序列需要差分多少次才能平稳。同时，相关分析图还可以识别季节性成分，甚至可识别那些由时间序列本身无法辨别的季节性峰值。因此，相关分析图可以说是分析随机时间序列特征的一种方便、有效的工具，希望同学们能够熟练掌握。

案例五　上证 A、B 股指数协整关系检验及误差修正模型

在本案例中，我们利用我国上海证券交易所 A、B 股指数 1998 年 1 月 9 日到 2008 年 3 月 7 日周收盘价数据[①]（参见资料光盘：数据集/单位根检验、协整与误差修正模型数据/上证 A、B 股指数周数据.xls），介绍单位根检验、协整关系检验的基本方法以及误差修正模型的应用。

1. 创建 Eviews 工作文件（Workfile）

从 Eviews 主选单中选“File/New/Workfile”，选择“Undated or irregular”选项，输入“Start date：1 End date：500”，方法如案例一中所介绍的。虽然是周数据，但是由于我国股票市场受五一、十一、春节等假期的影响，数据只能选择 Undated or irregular 类型。

2. 录入数据，并对序列进行初步分析

在 Workfile 窗口中选“Objects/New Object”，新建一个序列对象，命名为 SHA，用来保存上证 A 股指数周收盘价数据，并将数据导入。该序列的折线图如图 5—1

① 数据来源：大智慧软件。

所示。

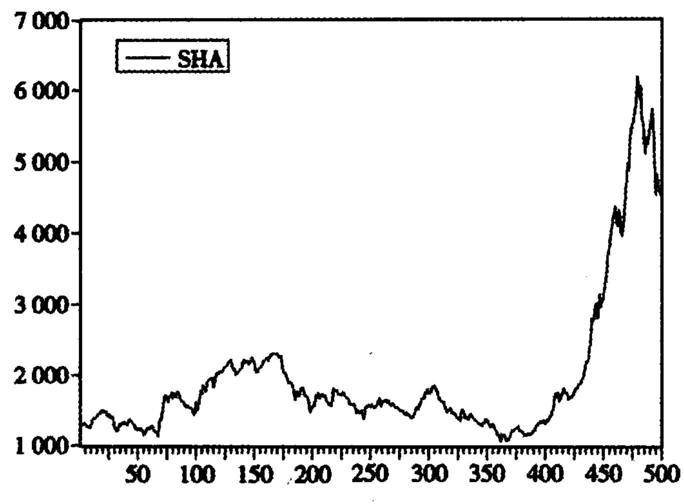


图 5—1　上证 A 股指数周收盘价序列的折线图

再新建一个序列对象，命名为 SHB，用来保存上证 B 股指数周收盘价数据，并将数据导入。该序列的折线图如图 5—2 所示。

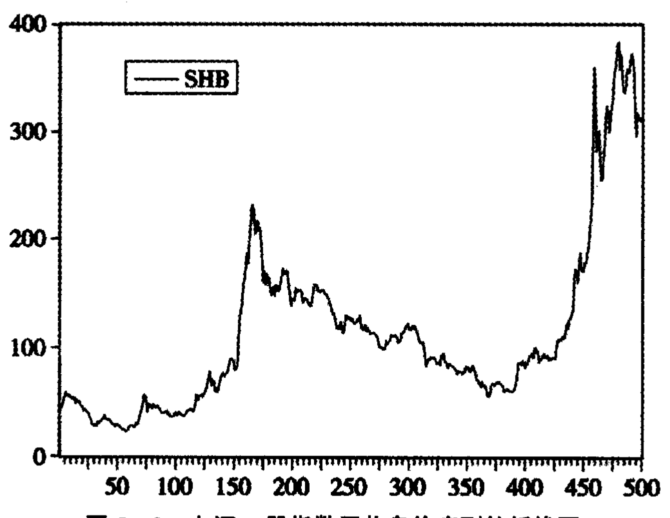


图 5—2　上证 B 股指数周收盘价序列的折线图

从图 5—1 和图 5—2 我们可以看到，上证 A、B 股指数周收盘价数据序列总体上来说有类似于随机游走过程的形式，都是非平稳的。为了更好地观察两者的关系，我们将两个序列以组的形式打开，并观察折线图，结果如图 5—3 所示。

从图 5—3 我们可以看到，上证 A、B 股指数周收盘价数据序列具有大致相同的趋势和变化规律，说明两者可能存在协整关系。

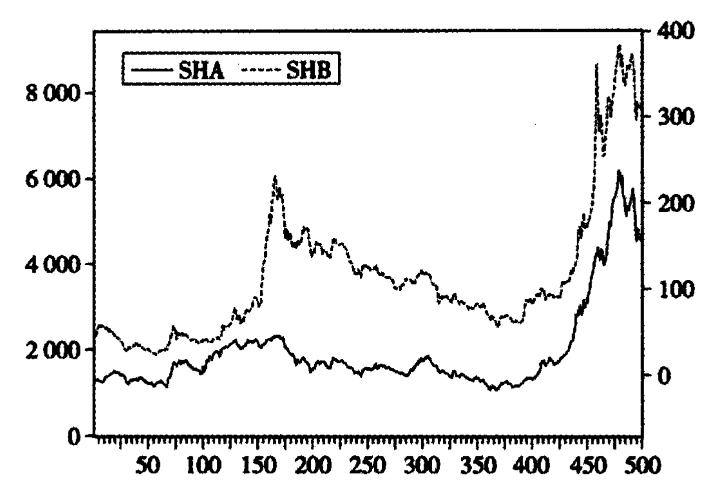


图 5—3 上证 A、B 股指数周收盘价序列的折线图

3. 单位根检验

为了避免虚假回归问题，我们首先对 SHA 和 SHB 序列进行单位根检验。

（1）SHA 序列单位根检验

为了确定 SHA 序列的非平稳性，首先我们观察相关分析图，方法如案例四中所介绍的，结果如图 5—4 和图 5—5 所示。图 5—4 是 SHA 原序列的相关分析图，图 5—5 是 SHA 一次差分序列的相关分析图。

从图 5—4 和图 5—5 我们看出，SHA 序列确实是非平稳的；SHA 序列一次差分后变得平稳，但是在 k=1，2，3 处，自相关系数显著不为零，这反映出高阶的序列相关性。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.990	0.990	493.40	0.000
		2	0.980	-0.057	977.26	0.000
		3	0.968	-0.051	1450.8	0.000
		4	0.955	-0.078	1912.6	0.000
		5	0.941	-0.044	2362.0	0.000
		6	0.926	0.049	2800.0	0.000
		7	0.913	-0.117	3224.6	0.000
		8	0.896	-0.107	3634.3	0.000
		9	0.878	-0.065	4028.0	0.000
		10	0.859	0.012	4406.1	0.000
		11	0.841	0.030	4768.9	0.000
		12	0.823	0.005	5117.0	0.000
		13	0.805	0.010	5461.0	0.000
		14	0.787	-0.039	5770.7	0.000
		15	0.769	0.049	6076.7	0.000

图 5—4 SHA 序列的相关分析图

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.108	0.108	5.8184	0.016
		2	0.121	0.111	13.206	0.001
		3	0.126	0.105	21.228	0.000
		4	0.048	0.014	22.390	0.000
		5	0.067	0.038	24.683	0.000
		6	-0.008	-0.037	24.714	0.000
		7	-0.009	-0.023	24.753	0.001
		8	-0.043	-0.050	25.715	0.001
		9	0.099	0.117	30.688	0.000
		10	0.051	0.047	32.000	0.000
		11	0.077	0.066	35.027	0.000
		12	0.003	-0.040	35.033	0.000
		13	0.038	0.014	35.758	0.001
		14	0.058	0.024	37.504	0.001
		15	-0.025	-0.041	37.827	0.001

图 5—5 SHA 序列一次差分后的相关分析图

下面我们使用更加准确的单位根检验方法对 SHA 序列的平稳性进行检验。

打开 SHA 序列，在序列窗口下选中“View/Unit Root Test”，如图 5—6 所示。

Series: SHA Workfile:

SpreadSheet
Line Graph
Bar Graph

Descriptive Statistics ▶
Tests for Descriptive Stats ▶
Distribution Graphs ▶
One-Way Tabulation...

Correlogram...
Unit Root Test...

Conversion Options...
Label

图 5—6 对 SHA 序列进行单位根检验

点击后出现如图 5—7 所示的单位根检验定义对话框。

其中，检验类型（Test Type）提供了(A)DF 检验和 PP 检验，这里我们选择(A)DF 检验。Test for unit root in 默认的 level 表示对原序列做单位根检验，1st difference 表示对一次差分序列做单位根检验，2nd difference 表示对二次差分序列做单位根检验。这里我们对原序列做单位根检验，选择“level”。Include in test equation 允许我们对检验方程的形式做出选择，Intercept 为加入常数项（位移项），适用于随机趋势过程；Trend and intercept 为加入趋势项和常数项，适用于趋势非平稳过程；None 为不加入

添加项，适用于随机游走过程。各种过程的特征可参考实验五中的实验准备知识部分的介绍。根据对图 5—1 的分析，我们应该选择“None”。Lags to include，为定义 ADF 检验时的最大滞后阶数 p，如果是 DF 检验可改为 0。从图 5—5 的分析，我们可以判断有高阶序列相关，应该使用 ADF 检验，对于 p 的选取我们使用 AIC 准则来确定，先从 p=1 开始。

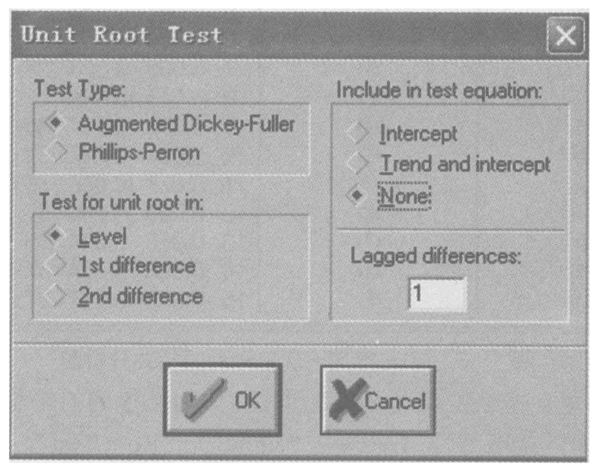


图 5—7 单位根检验对话框

点击“OK”后，得到如图 5—8 所示的结果。该输出结果上部显示 ADF 统计量值和 1%、5%、10%显著性水平下的临界值，下部显示检验方程的估计结果。

ADF Test Statistic	1.524781	1% Critical Value*	-2.5698
		5% Critical Value	-1.9401
		10% Critical Value	-1.6160

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(SHA)
Method: Least Squares
Date: 04/21/03 Time: 13:49
Sample(adjusted): 3 500
Included observations: 498 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
SHA(-1)	0.002615	0.001715	1.524781	0.1280
D(SHA(-1))	0.104711	0.044853	2.334558	0.0200

R-squared	0.011584	Mean dependent var	6.480044
Adjusted R-squared	0.009591	S.D. dependent var	84.51336
S.E. of regression	84.10711	[illegible]	[illegible]
Sum squared resid	3508707.	Schwarz criterion	11.72298
Log likelihood	-2912.811	Durbin-Watson stat	2.021394

图 5—8 SHA 序列单位根检验初步结果

此时（p=1），AIC 的值为 11.70607。为了最终确定 p 的取值，我们再分别取 p=2，3，…，直到 AIC 的值开始上升为止。将 AIC 的值汇总到表 5—1。

表 5—1　AIC 值汇总表

p	1	2	3	4
AIC 值	11.70607	11.70029	11.69539	11.70112

从表 5—1，我们看到 AIC 的值的最小值出现在 p=3 时，这也符合图 5—5 的分析。最终我们确定 p=3，再作 ADF 检验，结果如图 5—9 所示。

ADF Test Statistic	0.961082	1% Critical Value*	-2.5698
		5% Critical Value	-1.9401
		10% Critical Value	-1.6160

图 5—9　SHA 序列单位根检验最终结果

从图 5—9，我们看到 ADF 的值比 10%显著性水平下的临界值都大，不能拒绝原假设，说明 SHA 序列存在单位根，是非平稳的。

我们再对一次差分序列做单位根检验，与上面不同的是 Test for unit root in 选项选择“1st difference”，其他都相同。AIC 值的最小值出现在 p=2 时，因此 p 取 2，结果如图 5—10 所示。

ADF Test Statistic	-10.12007	1% Critical Value*	-2.5698
		5% Critical Value	-1.9401
		10% Critical Value	-1.6160

图 5—10　SHA 一次差分序列单位根检验结果

从图 5—10，我们看到 ADF 的值比 1%显著性水平下的临界值都小，所以拒绝原假设，说明 SHA 一次差分序列不存在单位根，是平稳的。也就是说，序列 SHA 为 1 阶单整序列，即 SHA~I（1）。

（2）SHB 序列单位根检验

该部分与上面 SHA 序列单位根检验基本一致，因此我们只作简单介绍。

首先观察相关分析图，图 5—11 是 SHB 原序列的相关分析图，图 5—12 是 SHB 一次差分序列的相关分析图。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.988	0.988	491.14	0.000
		2	0.974	-0.109	969.13	0.000
		3	0.960	0.034	1434.8	0.000
		4	0.947	0.004	1888.7	0.000
		5	0.933	-0.070	2329.7	0.000
		6	0.920	0.084	2759.5	0.000
		7	0.906	-0.076	3177.3	0.000
		8	0.890	-0.083	3581.7	0.000
		9	0.874	-0.024	3972.3	0.000
		10	0.857	-0.052	4348.6	0.000
		11	0.840	0.000	4710.7	0.000
		12	0.822	-0.062	5057.9	0.000
		13	0.802	-0.046	5389.8	0.000
		14	0.783	0.010	5706.8	0.000
		15	0.766	0.059	6010.6	0.000

图 5—11　SHB 序列的相关分析图

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.226	0.226	25.650	0.000
		2	-0.088	-0.146	29.532	0.000
		3	-0.021	0.037	29.758	0.000
		4	0.103	0.093	35.079	0.000
		5	-0.041	-0.099	35.927	0.000
		6	-0.058	0.000	37.644	0.000
		7	-0.040	-0.039	38.437	0.000
		8	-0.010	-0.013	38.488	0.000
		9	0.027	0.043	38.850	0.000
		10	0.055	0.040	40.410	0.000
		11	0.167	0.167	54.646	0.000
		12	0.080	0.006	57.919	0.000
		13	-0.051	-0.054	59.252	0.000
		14	-0.079	-0.049	62.480	0.000
		15	-0.028	-0.045	62.895	0.000

图 5—12 SHB 序列一次差分后的相关分析图

从图 5—11 和图 5—12 我们看出，SHA 序列确实是非平稳的；SHA 序列一次差分后变得平稳，但是在 k=1 处，自相关系数显著不为零，这反映出可能存在一定的序列相关性。

通过对图 5—2 的分析，我们做 Test for unit root in 选择“level”，Include in test equation 选择“None”的(A)DF 检验。*AIC* 值的最小值出现在 p=2 时，这也符合图 5—12 的分析。最终我们确定 p=2，做 ADF 检验，结果如图 5—13 所示。

ADF Test Statistic	0.993170	1% Critical Value*	-2.5698
		5% Critical Value	-1.9401
		10% Critical Value	-1.6160

图 5—13 SHB 序列单位根检验最终结果

从图 5—13，我们看到 ADF 的值比 10%显著性水平下的临界值都大，不能拒绝原假设，说明 SHB 序列存在单位根，是非平稳的。

我们再对 SHB 一次差分序列做单位根检验，Test for unit root in 选择“1st difference”，*AIC* 的值的最小值出现在 p=1，因此 p 取 1，结果如图 5—14 所示。

ADF Test Statistic	-15.95455	1% Critical Value*	-2.5698
		5% Critical Value	-1.9401
		10% Critical Value	-1.6160

图 5—14 SHB 一次差分序列单位根检验结果

从图 5—14，我们看到 ADF 的值比 1%显著性水平下的临界值都小，所以拒绝原假设，说明 SHB 一次差分序列不存在单位根，是平稳的。也就是说，序列 SHB 为 1 阶单整序列，即 SHB~I（1）。

4. 协整检验

经过上面的单位根检验，我们得出结论 SHA 和 SHB 序列都是 1 阶单整序列，满足协整关系检验的前提，即协整关系检验的的第一步已经完成。

第二步，用变量 SHB 对 SHA 进行普通最小二乘回归。建立方程，估计方法选择最小二乘（LS），在方程定义对话窗输入 shb c sha，得到如图 5—15 所示的估计结果。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-19.08139	3.523489	-5.415480	0.0000
SHA	0.067831	0.001588	42.72059	0.0000

R-squared	0.785627	Mean dependent var	113.6449
Adjusted R-squared	0.785196	S.D. dependent var	80.18888
S.E. of regression	37.16510	Akaike info criterion	10.07261
Sum squared resid	687859.7	Schwarz criterion	10.08947
Log likelihood	-2516.152	F-statistic	1825.049
Durbin-Watson stat	0.034183	Prob(F-statistic)	0.000000

图 5—15　协整方程输出结果

估计方程为：

$$SHB = -19.081 + 0.068SHA \qquad (5.1)$$

$$t=（-5.415）（42.72） \qquad R^2 = 0.786, DW = 0.034$$

为了方便对残差项进行单位根检验，我们将残差项保存在序列 u 中，方法是：在方程窗口中选中“Procs/Make Residual Series”，如图 5—16 所示。点击后，将残差项起名为 u，如图 5—17 所示。

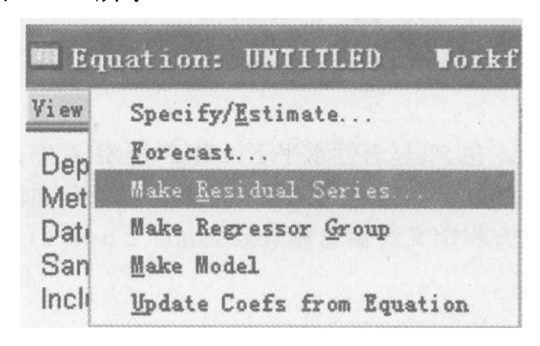


图 5—16　保存残差项序列

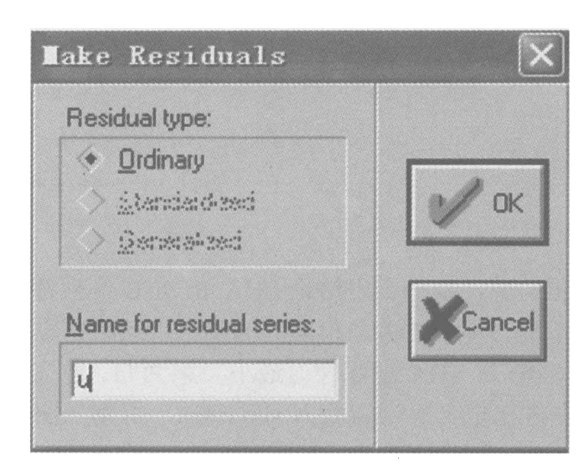


图 5—17 保存残差项到序列 u 中

接下来对残差序列 u 作单位根检验。Test for unit root in 选择 “level”，Include in test equation 选择 “None”。*AIC* 的值的最小值出现在 *p*=2 时，此时 ADF 检验结果如图 5—18 所示。

ADF Test Statistic	-1.988142	1% Critical Value*	-2.5698
		5% Critical Value	-1.9401
		10% Critical Value	-1.6160

图 5—18 SHB 序列单位根检验最终结果

从图 5—18，我们看到 ADF 的值比 5%显著性水平下的临界值小，比 1%显著性水平下的临界值大，因此，我们可以说在 5%显著性水平下拒绝原假设，说明残差序列 u 不存在单位根，是平稳的。也就是说，序列 u 为 0 阶单整序列，即 u~*I*（0），SHB 和 SHA 序列存在协整关系，方程（5.1）即协整方程，协整向量为（1，-0.068）。如果显著性水平取 1%，则结论变为 SHB 和 SHA 序列不存在协整关系。

5. 建立误差修正模型（ECM）

上面我们得到结论，在 5%显著性水平下，SHB 和 SHA 序列存在协整关系。可以建立误差修正模型（ECM）。误差修正项 *ecm*，就是我们前面协整方程得到的残差序列 u。建立方程，在方程定义对话窗输入 d（shb）c u（-1）d（sha），得到如图 5—19 所示的估计结果。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.188996	0.302604	0.624566	0.5325
U(-1)	-0.018855	0.008140	-2.316408	0.0209
D(SHA)	0.053068	0.003581	14.81832	0.0000
R-squared	0.315792	Mean dependent var	0.530080	
Adjusted R-squared	0.313034	S.D. dependent var	8.132136	
S.E. of regression	6.740194	Akaike info criterion	6.660048	
Sum squared resid	22533.39	Schwarz criterion	6.685375	
Log likelihood	-1658.682	F-statistic	114.4631	
Durbin-Watson stat	1.721728	Prob(F-statistic)	0.000000	

图 5—19　误差修正模型输出结果

但是常数项不显著，因此重新定义方程，输入 d（shb）u（-1）d（sha），得到如图 5—20 所示的估计结果。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
U(-1)	-0.018842	0.008135	-2.316240	0.0210
D(SHA)	0.053238	0.003569	14.91774	0.0000
R-squared	0.315254	Mean dependent var	0.530080	
Adjusted R-squared	0.313877	S.D. dependent var	8.132136	
S.E. of regression	6.736057	Akaike info criterion	6.656827	
Sum squared resid	22551.11	Schwarz criterion	6.673711	
Log likelihood	-1658.878	Durbin-Watson stat	1.721283	

图 5—20　修改后的误差修正模型输出结果

可知误差修正模型形式为：

$$\Delta SHB = -0.0189ecm_{t-1} + 0.053\Delta SHA + v_t \qquad (5.2)$$

（-2.316）　　（14.918）　　　R^2=0.315，DW=1.721

根据协整方程（5.1）式，可知误差修正项为：

$$ecm = SHB + 19.081 - 0.068SHA \qquad (5.3)$$

误差修正模型（5.2）反映了上证 B 股指数周收盘价的短期波动如何被决定。上证 B 股指数周收盘价的短期波动可以分为两个部分：一部分是上证 A 股指数周收盘价短期波动的影响，另一部分是偏离长期均衡的影响。误差修正项的系数-0.0189 反映了对偏离长期均衡的调整力度，当短期波动偏离长期均衡时，将以-0.0189 的调整力度将非均衡状态拉回到均衡状态。

为了避免虚假回归问题，通常采用差分的方法使序列平稳之后再建立模型。但为了保证模型的经济意义，我们一般先做协整检验，再建立模型。因此，单位根检验、协整关系检验已经成为数据分析的基础性工作，希望同学们能够熟练掌握。

案例六 美元对欧元汇率 ARMA 模型应用

在本案例中，我们利用美元对欧元汇率 1993 年 1 月到 2007 年 12 月的月均价数据[①]（参见资料光盘：数据集/ARMA 模型应用数据/美元对欧元汇率月均价数据.xls），介绍 ARIMA 模型的识别、估计、检验及预测的方法。

1. 创建 Eviews 工作文件（Workfile）

从 Eviews 主选单中选“File/New/Workfile”，选择“Monthly”选项，输入“Start date：1993:01 End date：2007:12”，方法如案例一中所介绍的。

2. 录入数据，并对序列进行初步分析

在 Workfile 窗口中选“Objects/New Object”，新建一个序列对象，命名为 EURO，用来保存美元对欧元汇率的月均价数据，并将数据导入。该序列的折线图如图 6—1 所示。

从图 6—1 我们可以看到，美元对欧元汇率在 2001 年左右处于高位，2002 年以

① 数据来源：The University of British Columbia Sauder School of Business Pacific Exchange Rate Service,网址 http://fx.sauder.ubc.ca/。

后一直处于下跌的态势。数据序列总体上来说有类似于随机游走过程的形式，应该是非平稳的。



图 6—1　美元对欧元汇率月均价序列的折线图

3. ARIMA（p，d，q）模型阶数识别

（1）确定单整阶数 d

我们运用前面介绍的单位根检验来确定单整阶数 d。首先对原序列进行 ADF 检验，检验方程选择没有添加项，根据 AIC 准则确定 ADF 检验的滞后阶数 p=2（不是 ARIMA（p，d，q）模型的自回归阶数），检验结果如图 6—2 所示。

ADF Test Statistic	-0.593799	1% Critical Value*	-2.5772
		5% Critical Value	-1.9415
		10% Critical Value	-1.6166

图 6—2　EURO 序列单位根检验最终结果

从图 6—2，我们看到 ADF 的值比 10%显著性水平下的临界值都大，不能拒绝原假设，说明 EURO 序列存在单位根，是非平稳的。

我们再对 EURO 序列的一次差分序列做单位根检验，根据 AIC 准则确定 ADF 检验的滞后阶数 p=1，检验结果如图 6—3 所示。

ADF Test Statistic	-9.540143	1% Critical Value*	-2.5772
		5% Critical Value	-1.9415
		10% Critical Value	-1.6166

图 6—3　EURO 一次差分序列单位根检验结果

从图 6—3，我们看到 ADF 的值比 1%显著性水平下的临界值都小，所以拒绝原假设，说明 EURO 一次差分序列不存在单位根，是平稳的。也就是说，序列 EURO 为 1 阶单整序列，即 EURO~*I*（1）。所以，可以确定单整阶数 *d*=1。

（2）确定自回归阶数 *p* 和移动平均阶数 *q*

创建 EURO 的一次差分序列，命名为 DEURO，下面利用 DEURO 序列的相关分析图来确定自回归阶数 *p* 和移动平均阶数 *q*。

图 6—4 是 DEURO 序列的相关分析图。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.308	0.308	17.279	0.000
		2	-0.087	-0.200	18.649	0.000
		3	-0.115	-0.024	21.071	0.000
		4	-0.018	0.017	21.129	0.000
		5	0.076	0.062	22.218	0.000
		6	0.039	-0.016	22.504	0.001
		7	0.072	0.095	23.493	0.001
		8	0.127	0.100	26.569	0.001
		9	0.070	0.014	27.511	0.001
		10	0.117	0.144	30.136	0.001
		11	0.068	0.015	31.017	0.001
		12	-0.016	-0.020	31.068	0.002
		13	-0.022	0.007	31.160	0.003
		14	-0.028	-0.033	31.309	0.005
		15	0.007	-0.016	31.318	0.008

图 6—4 DEURO 序列的相关分析图

从图 6—4 可以看出，DEURO 序列的偏自相关函数在滞后 1 期和滞后 2 期处显著不为零，可初步判定自回归阶数 *p*=2；自相关函数在滞后 1 期处显著不为零，初步判定移动平均阶数 *q*=1。至于 *p* 和 *q* 的最终确定，还要从低阶开始逐步试探，直到定出合适的模型为止。故初步选择适合 EURO 序列的模型可以有以下几个：ARIMA（1，1，0）、ARIMA（2，1，0）、ARIMA（0，1，1）、ARIMA（1，1，1）和 ARIMA（2，1，1）。

4. ARIMA（*p*，*d*，*q*）模型估计与检验

（1）ARIMA（1，1，0）模型估计与检验

建立 ARIMA（1，1，0）模型。建立方程，估计方法选择“LS-Least Squares（NLS

and ARMA)”，在方程定义对话窗输入“d（euro）c ar（1)”，得到如图6—5所示的估计结果。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000922	0.002097	-0.439890	0.6606
AR(1)	0.308279	0.071484	4.312552	0.0000
R-squared	0.095572	Mean dependent var		-0.000886
Adjusted R-squared	0.090433	S.D. dependent var		0.020291
S.E. of regression	0.019352	Akaike info criterion		-5.040911
Sum squared resid	0.065909	Schwarz criterion		-5.005160
Log likelihood	450.6411	F-statistic		18.59810
Durbin-Watson stat	1.871573	Prob(F-statistic)		0.000027
Inverted AR Roots	.31			

图6—5 ARIMA（1，1，0）模型初步估计输出结果

可见，常数项不显著，重新估计，方程定义对话窗输入“d（euro）ar（1)”，得到如图6—6所示的估计结果。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.309522	0.071265	4.343228	0.0000
R-squared	0.094579	Mean dependent var		-0.000886
Adjusted R-squared	0.094579	S.D. dependent var		0.020291
S.E. of regression	0.019307	Akaike info criterion		-5.051050
Sum squared resid	0.065982	Schwarz criterion		-5.033174
Log likelihood	450.5434	Durbin-Watson stat		1.871488
Inverted AR Roots	.31			

图6—6 ARIMA（1，1，0）模型最终估计输出结果

其中，Inverted AR Roots（自回归特征方程根的倒数）是0.31，在单位圆之内，说明模型是平稳的。

我们还得到几个有用的检验结果：Adjusted R^2=0.095，AIC=-5.05，SC=-5.03。

对残差做白噪声检验。在方程窗口中选中“View/Residual Tests/Correlogram – Q-statistics”，如图6—7所示。

点击后，出现如图6—8所示的对话框，要求填入最大滞后期，根据实验四中的介绍，该残差序列样本容量为178，我们可以取[178/10]或者[$\sqrt{178}$]，这里我们取13。

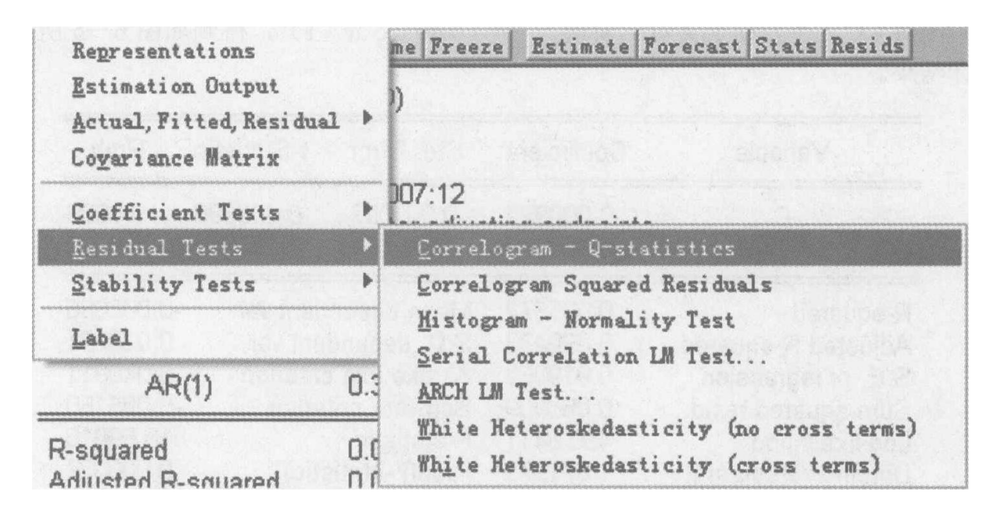


图 6—7　对残差做白噪声检验



图 6—8　对残差做白噪声检验对话框

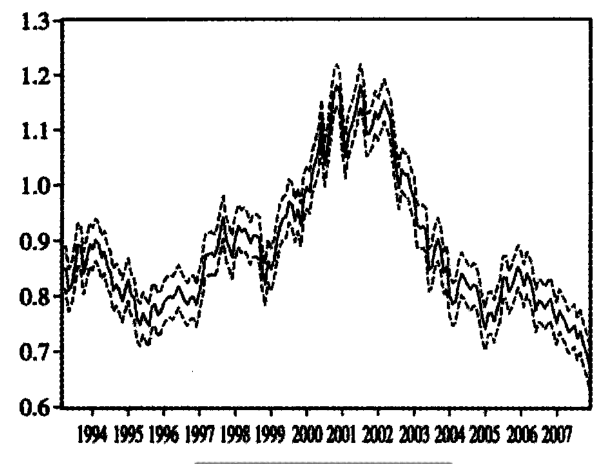
点击“OK”，得到如图 6—9 所示的结果。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.062	0.062	0.6943	
		2	-0.163	-0.167	5.5318	0.019
		3	-0.108	-0.089	7.6630	0.022
		4	-0.017	-0.033	7.7151	0.052
		5	0.077	0.050	8.8109	0.066
		6	-0.003	-0.029	8.8125	0.117
		7	0.043	0.063	9.1607	0.165
		8	0.099	0.104	11.021	0.138
		9	-0.005	0.001	11.027	0.200
		10	0.097	0.143	12.829	0.171
		11	0.043	0.058	13.189	0.213
		12	-0.033	-0.001	13.399	0.268
		13	-0.006	0.024	13.406	0.340

图 6—9　残差序列白噪声检验结果

从 K=13 一行找到 Q 统计量的值为 13.406，相伴概率（记为 p-Q）为 0.340>0.05，因此不能拒绝序列相互独立的原假设，残差序列为白噪声的检验通过。

利用该模型做静态预测，得到如图 6—10 所示的预测结果，MAPE=1.745。图 6—11 为 ARIMA（1，1，0）模型预测值与真实值的比较图。



Forecast: EUROF
Actual: EURO
Forecast sample: 1993:01 2007:
Adjusted sample: 1993:03 2007:
Included observations: 178

Root Mean Squared Error 0.019253
Mean Absolute Error 0.015580
Mean Abs. Percent Error 1.744810
Theil Inequality Coefficient 0.010770
Bias Proportion 0.001095
Variance Proportion 0.000020
Covariance Proportion 0.998885

图 6—10　ARIMA（1，1，0）模型预测结果



图 6—11　ARIMA（1，1，0）模型预测值与真实值比较图

通过 EURO 序列和 EUROF 序列观察 2007 年 12 月的真实值与预测值，并比较，如图 6—12 所示。

obs	EURO	EUROF
2007:12	0.686860	0.674492

图 6—12　2007 年 12 月的真实值与预测值

预测误差为：$\eta = \frac{0.674492 - 0.686860}{0.686860} = -0.018$

（2）其他几个模型估计与检验

类似 ARIMA（1，1，0）模型的操作，建立 ARIMA（2，1，0）、ARIMA（0，1，1）、ARIMA（1，1，1）、ARIMA（2，1，1）模型，但 ARIMA（1，1，1）、ARIMA（2，1，1）模型没有通过检验。具体估计与检验过程略。

（3）模型的评价与选择

下面将 ARIMA（1，1，0）、ARIMA（2，1，0）、ARIMA（0，1，1）模型的参数估计结果和相关检验结果汇总到表 6—1 和表 6—2。

表 6—1 **各模型参数估计结果汇总表**

模型	ϕ_1	ϕ_2	θ_1
ARIMA（1，1，0）	0.310	—	—
ARIMA（2，1，0）	0.373	-0.202	—
ARIMA（0，1，1）	—	—	0.385

表 6—2 各模型相关检验结果指标汇总表

模型	Adjusted R^2	AIC	SC	p-Q	MAPE	η
ARIMA（1，1，0）	0.095	-5.051	-5.033	0.340	1.745	-0.018
ARIMA（2，1，0）	0.126	-5.076	-5.040	0.698	1.696	-0.015
ARIMA（0，1，1）	0.122	-5.082	-5.064	0.727	1.699	-0.019

在三个模型都通过了参数显著性水平检验、模型平稳性和可逆性检验以及残差序列白噪声检验的前提下，我们通过表 6—2 来选择最优模型。6 个指标中，Adjusted R^2 和 p-Q 应该是越大越好，其他 4 个是越小越好。可以看出 ARIMA（1，1，0）模型明显不好。但是 ARIMA（2，1，0）和 ARIMA（0，1，1）模型难分伯仲。ARIMA（0，1，1）模型在模型的简洁性角度明显占优，而 ARIMA（2，1，0）模型在预测方面明显更好。出于预测的目的，我们最终选择 ARIMA（2，1，0）模型。

5. ARIMA（p，d，q）模型外推预测

从上面的分析，我们最终选择 ARIMA（2，1，0）模型作为最优模型，下面我们利用 ARIMA（2，1，0）模型对 2008 年 1 月美元对欧元汇率的月均价进行外推预测。方法参看案例二。

预测结果如图 6—13 所示，其中 0.693 为 2008 年 1 月美元对欧元汇率的月均价的预测值，0.655～0.732 为加减 2 倍标准误之后的范围。

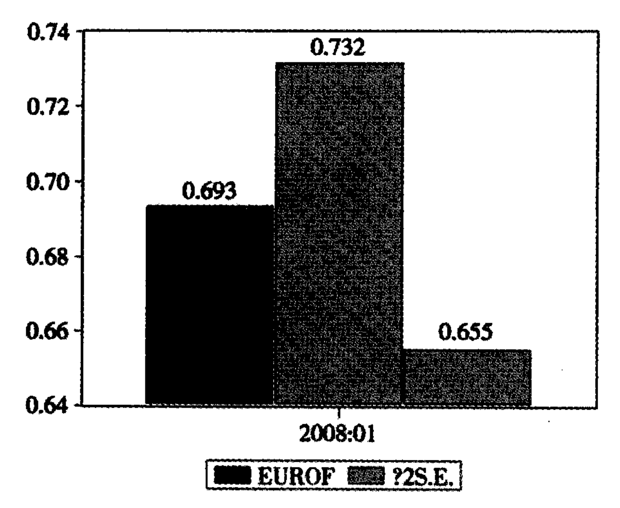


图6—13　2008年1月美元对欧元汇率月均价预测值

案例七　上证 A 股收益率条件异方差模型应用

在本案例中，我们利用我国上海证券交易所 A 股指数 1998 年 1 月 9 日到 2008 年 3 月 7 日周收盘价数据[①]（参见资料光盘：数据集/条件异方差模型应用数据/上证 A 股指数周数据.xls），介绍条件异方差模型应用的基本方法。

1. 创建 Eviews 工作文件（Workfile）

从 Eviews 主选单中选“File/New/Workfile”，选择“Undated or irregular”选项，输入“Start date：1 End date：500”，方法如案例一中所介绍的。

2. 录入数据，并对序列进行初步分析

在 Workfile 窗口中选“Objects/New Object”，新建一个序列对象，命名为 SHA，用来保存上证 A 股指数周收盘价数据，并将数据导入。该序列的折线图如图 7—1 所示。

对原数据进行对数处理，新建一个序列对象，命名为 LNSHA，用来保存上证 A 股指数周收盘价数据的对数序列，并用公式 LNSHA=log（SHA）生成。LNSHA 序

① 数据来源：大智慧软件。

列的折线图如图 7—2 所示。

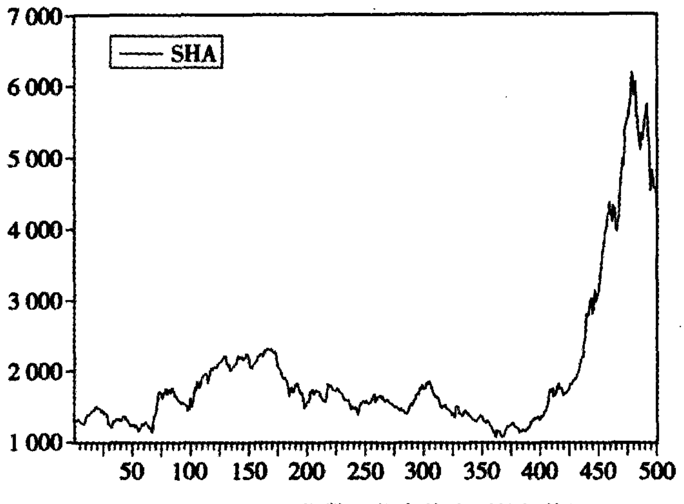


图 7—1　上证 A 股指数周收盘价序列的折线图

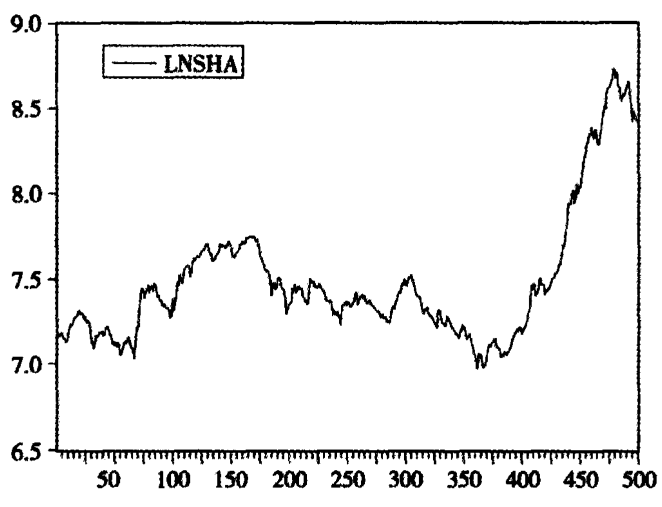


图 7—2　上证 A 股指数周收盘价对数序列的折线图

从图 7—1 和图 7—2 我们可以看到，上证 A 股指数周收盘价数据原序列和对数序列总体上来说有类似于随机游走过程的形式，这与我们学过的金融学知识相符合。

3. 建立主体模型

从图 7—1 和图 7—2 的分析和金融学知识，我们初步选定一阶自回归模型作为

主体模型，模型形式为 $y_t = \phi_1 y_{t-1} + \delta + u_t$。新建方程，在方程定义对话窗中输入模型形式：log（SHA）c log（SHA（-1））（为了预测方便，这里采用函数形式）。估计结果如图 7—3 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.009796	0.028328	-0.345825	0.7296
LOG(SHA(-1))	1.001639	0.003779	265.0252	0.0000
R-squared	0.992974	Mean dependent var		7.487855
Adjusted R-squared	0.992960	S.D. dependent var		0.387205
S.E. of regression	0.032489	Akaike info criterion		-4.011824
Sum squared resid	0.524604	Schwarz criterion		-3.994940
Log likelihood	1002.950	F-statistic		70238.33
Durbin-Watson stat	1.839442	Prob(F-statistic)		0.000000

图 7—3 一阶自回归模型方程输出结果

常数项没有通过显著性检验，因此从方程中剔除，重新定义方程为 log（SHA）log（SHA（-1））。得到估计结果如图 7—4 所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(SHA(-1))	1.000334	0.000194	5159.753	0.0000
R-squared	0.992972	Mean dependent var		7.487855
Adjusted R-squared	0.992972	S.D. dependent var		0.387205
S.E. of regression	0.032460	Akaike info criterion		-4.015592
Sum squared resid	0.524730	Schwarz criterion		-4.007149
Log likelihood	1002.890	Durbin-Watson stat		1.836600

图 7—4 调整后的一阶自回归模型方程输出结果

从估计结果来看，系数通过显著性检验，拟合优度达到 0.993，初步看来效果不错。我们再来看看残差图，如图 7—5 所示，为了看出时间，我们在横轴加入时间标志（以后各图形不再加时间标志，请读者参考本图）。

从图 7—5 我们可以看到，残差的波动有聚类的现象，波动在一些时间内比较小（比如 2001 年、2004 年左右），在一些时间内比较大（比如 2000 年、2007 年左右）。我们再看看残差平方序列图，如图 7—6 所示，也出现了聚类现象。这些说明误差项可能具有条件异方差性（即 ARCH 效应）。

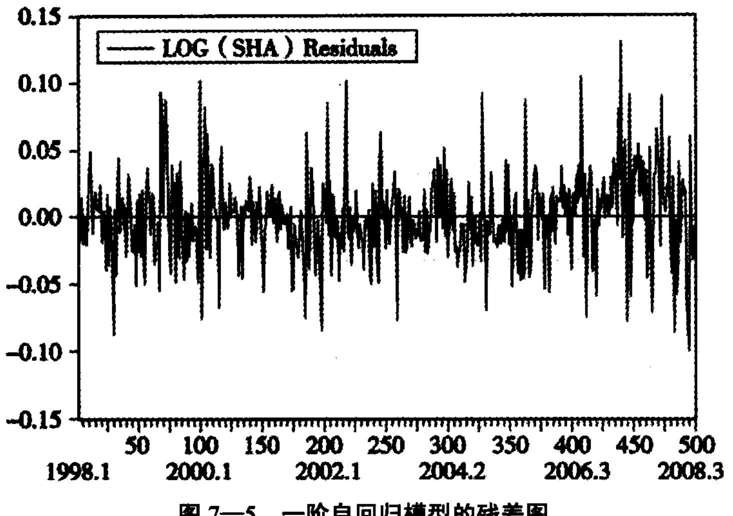


图7—5 一阶自回归模型的残差图

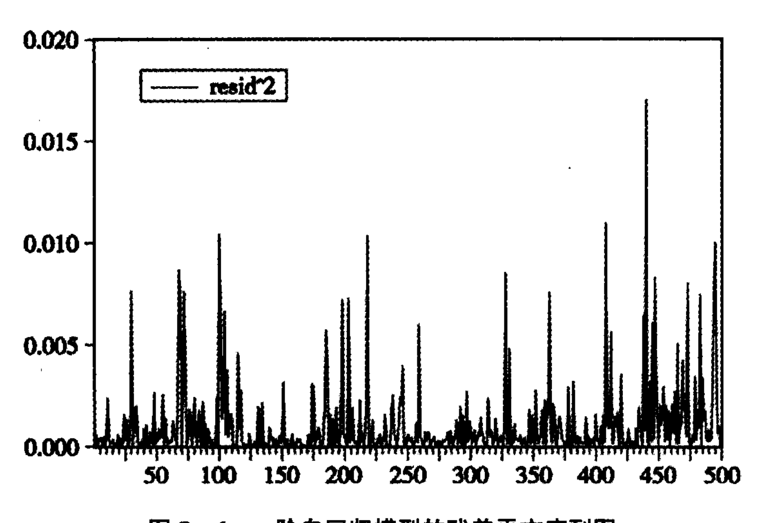


图7—6 一阶自回归模型的残差平方序列图

4. ARCH 效应检验

（1）ARCH-LM 检验

为了检验 ARCH 效应，我们利用 ARCH-LM 检验。在方程窗口中选中“View/Residual Tests/ARCH LM Test”，如图 7—7 所示。

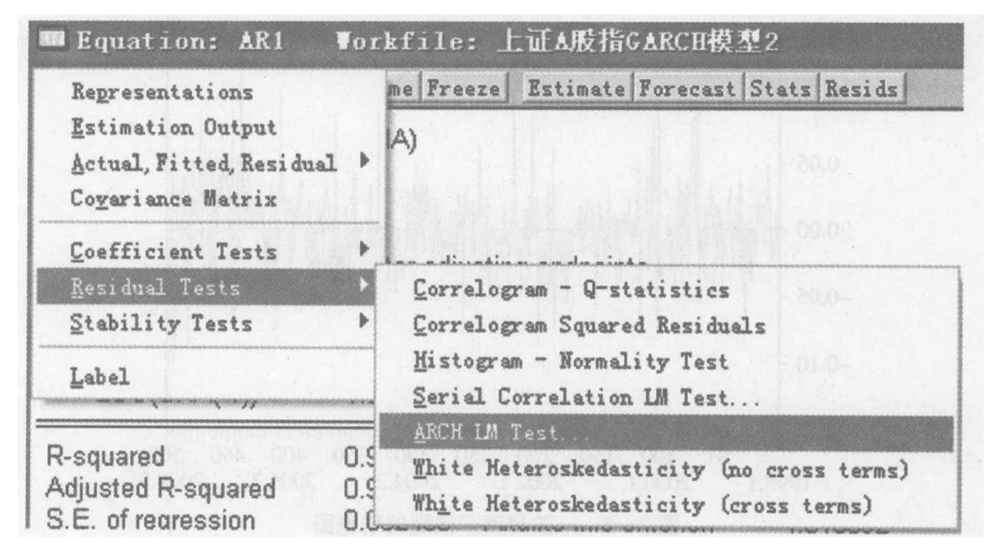


图 7—7 进行 ARCH-LM 检验

点击后出现如图 7—8 所示的对话框。要求填入包含的滞后阶数，系统默认为 1，这里我们先不作改变。

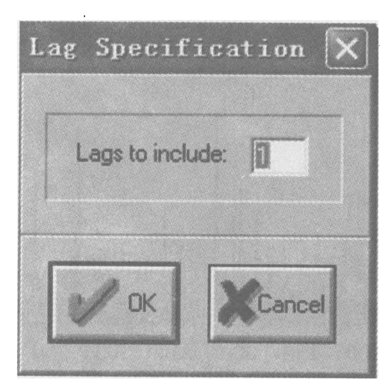


图 7—8 ARCH-LM 检验对话框

点击“OK”后，我们得到 ARCH-LM 检验结果，如图 7—9 所示。

ARCH Test:

F-statistic	7.990433	Probability	0.004892
Obs*R-squared	7.895458	Probability	0.004956

图 7—9 ARCH-LM 检验结果

图7—9中第二行Obs*R-squared，即ARCH-LM统计量。可以看到该统计量的相伴概率为0.005，小于0.05，拒绝没有ARCH效应的原假设，说明残差序列存在ARCH效应。

（2）残差平方序列的相关分析图和残差平方的 Q 统计量检验

我们还可以通过残差平方序列的相关分析图来检验 ARCH 效应。在方程窗口中选中“View/Residual Tests/Correlogram Squared Residuals”，如图 7—10 所示。

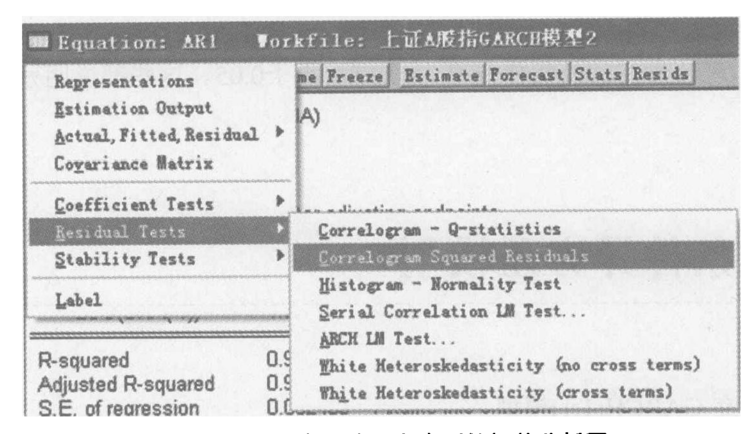


图 7—10 观察残差平方序列的相关分析图

点击后出现如图 7—11 所示的对话框。要求填入包含的滞后阶数，由于这里我们还要做残差平方的 Q 统计量检验，因此取样本数的平方根 22。



图 7—11 残差平方序列相关分析图对话框

点击“OK”后，得到残差平方序列的相关分析图，如图 7—12 所示。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.126	0.126	7.9547	0.005
		2	0.139	0.125	17.645	0.000
		3	0.057	0.027	19.281	0.000
		4	0.125	0.102	27.229	0.000
		5	0.054	0.020	28.700	0.000
		6	0.008	-0.030	28.729	0.000
		7	0.027	0.014	29.100	0.000
		8	0.048	0.033	30.258	0.000
		9	0.016	-0.004	30.366	0.000
		10	-0.011	-0.022	30.449	0.001
		11	0.020	0.018	30.654	0.001
		12	0.039	0.031	31.434	0.002
		13	0.017	0.004	31.586	0.003
		14	0.057	0.053	33.279	0.003
		15	0.031	0.013	33.760	0.004
		16	-0.009	-0.039	33.802	0.006
		17	-0.044	-0.053	34.811	0.007
		18	0.007	0.012	34.833	0.010
		19	-0.038	-0.039	35.589	0.012
		20	0.077	0.092	38.659	0.007
		21	-0.010	-0.007	38.711	0.011
		22	-0.009	-0.029	38.752	0.015

图 7—12 残差平方序列的相关分析图

图7—12可以看到，残差平方序列的自相关系数在1、2、4处显著不为0。同时，残差平方的Q统计量在22处相伴概率为0.015，小于0.05。这些都说明残差序列存在ARCH效应。

5. 建立条件异方差模型

（1）建立 GARCH(1,1)模型

从上面的分析，我们认为一阶自回归模型的残差项具有 ARCH 效应，因此我们利用最常用的 GARCH(1,1)模型重新建模。

从 Eviews 主选单中选“Quick/Estimate Equation”，或者用新建对象的方法新建一个 Equation 对象，在估计方法下拉选单中选择 ARCH 模型，如图 7—13 所示。

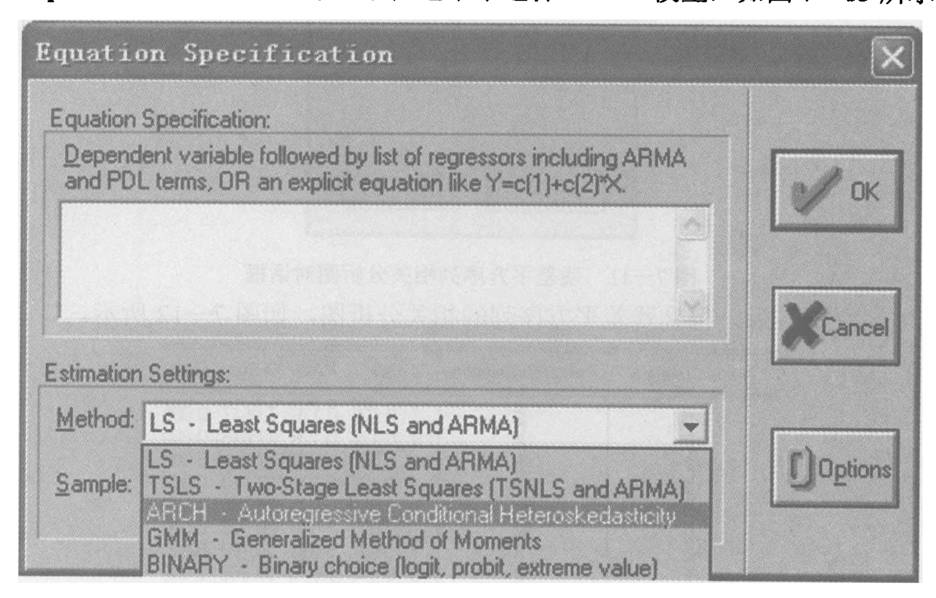


图 7—13　建立条件异方差模型

点击后出现如图 7—14 所示的对话框。

在 Mean Equation Specification 框中输入均值方程（即主体模型），本例中输入“log（SHA）log（SHA（-1））”。ARCH-M term 为 ARCH-M 模型或 GARCH-M 模型选项，即条件方差 h_t（或者标准差 $\sqrt{h_t}$ ）是否加入均值方程中，这里先不加入，选择“None”。ARCH Specification 为条件方差方程定义 GARCH（p，q）的阶数，如果是 ARCH(q)模型，将 GARCH 后的 1 改为 0，本例为 GARCH(1,1)模型，可以保持默认值不变。阶数定义框下面的几个按钮为各种 GARCH 模型的变形，本例中选

择 GARCH 模型。点击“OK”后，得到 GARCH(1,1)模型估计结果（如图 7—15 所示）。

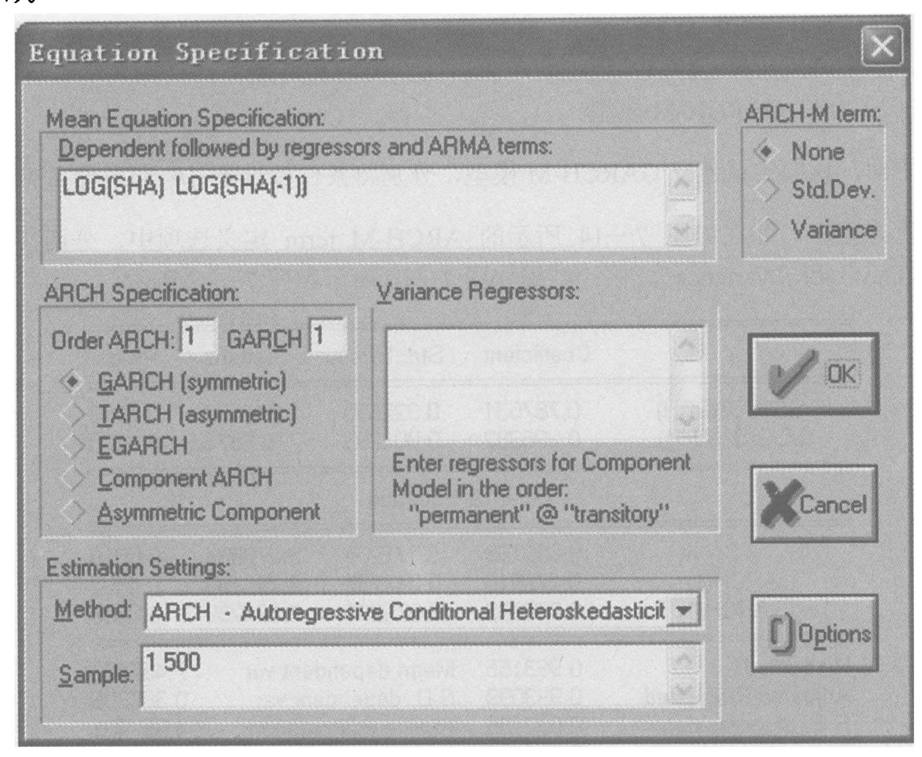


图 7—14 建立条件异方差模型对话框

	Coefficient	Std. Error	z-Statistic	Prob.
LOG(SHA(-1))	1.000080	0.000173	5777.781	0.0000
		Variance Equation		
C	0.000185	6.14E-05	3.007500	0.0026
ARCH(1)	0.206859	0.051804	3.993110	0.0001
GARCH(1)	0.631443	0.085533	7.382473	0.0000
R-squared	0.992948	Mean dependent var	7.487855	
Adjusted R-squared	0.992905	S.D. dependent var	0.387205	
S.E. of regression	0.032615	Akaike info criterion	-4.076548	
Sum squared resid	0.526537	Schwarz criterion	-4.042780	
Log likelihood	1021.099	Durbin-Watson stat	1.829834	

图 7—15 GARCH（1，1）模型输出结果

输出结果上部为均值方程的系数估计结果，中间为条件方差方程的系数估计结果，下部为各种指标的报告。从图 7—15 可以看到，各参数通过显著性检验。条件

方差方程的系数$(\alpha_1+\beta_1)=0.838<1$，满足参数约束条件，说明条件方差平稳。与图7—4的一阶自回归模型结果比较，虽然拟合优度并没有改善，但是*AIC*和*SC*明显变小，说明GARCH（1，1）模型对一阶自回归模型有所改善。

（2）建立GARCH-M模型

下面，我们尝试建立GARCH-M模型，分别将条件标准差$\sqrt{h_t}$和条件方差h_t加入均值方程中，即在如图7—14所示的ARCH-M term定义选项中，分两次选择“Std.Dev”和“Variance”。估计结果如图7—16所示和图7—17所示。

	Coefficient	Std. Error	z-Statistic	Prob.
SQR(GARCH)	0.787631	0.322590	2.441588	0.0146
LOG(SHA(-1))	0.996882	0.001293	770.6876	0.0000
	Variance Equation			
C	0.000169	5.14E-05	3.278934	0.0010
ARCH(1)	0.176910	0.047237	3.745186	0.0002
GARCH(1)	0.669663	0.075291	8.894299	0.0000
R-squared	0.993155	Mean dependent var	7.487855	
Adjusted R-squared	0.993099	S.D. dependent var	0.387205	
S.E. of regression	0.032165	Akaike info criterion	-4.087281	
Sum squared resid	0.511100	Schwarz criterion	-4.045071	
Log likelihood	1024.777	Durbin-Watson stat	1.899844	

图7—16　GARCH-M模型（加入条件标准差）输出结果

	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	9.825256	4.509169	2.178950	0.0293
LOG(SHA(-1))	0.998843	0.000561	1780.988	0.0000
	Variance Equation			
C	0.000167	5.39E-05	3.092929	0.0020
ARCH(1)	0.175675	0.046954	3.741459	0.0002
GARCH(1)	0.672977	0.077746	8.656138	0.0000
R-squared	0.993125	Mean dependent var	7.487855	
Adjusted R-squared	0.993069	S.D. dependent var	0.387205	
S.E. of regression	0.032235	Akaike info criterion	-4.084439	
Sum squared resid	0.513320	Schwarz criterion	-4.042229	
Log likelihood	1024.068	Durbin-Watson stat	1.899835	

图7—17　GARCH-M模型（加入条件方差）输出结果

两个GARCH-M模型比较，加入条件标准差$\sqrt{h_t}$的GARCH-M模型优于加入条件方差h_t的GARCH-M模型。再比较加入条件标准差$\sqrt{h_t}$的GARCH-M模型与GARCH（1，1）模型，GARCH-M模型的拟合优度提高，*AIC*和*SC*减小，说明GARCH-M模型更好。

最终，我们认为加入条件标准差$\sqrt{h_t}$的GARCH-M模型是最优模型，该模型的估计结果为：

$$\log(SHA)_t = 0.9969\log(SHA)_{t-1} + 0.7876\sqrt{h_t}$$

（770.688）　　（2.442）

$$h_t = 0.00017 + 0.1769u_{t-1}^2 + 0.6697h_{t-1}$$

（3.789）　（3.745）　（8.894）　　$R^2 = 0.993$

各参数通过显著性检验。条件方差方程的系数$(\alpha_1+\beta_1)=0.8466<1$，满足参数约束条件。同时，$(\alpha_1+\beta_1)$作为衰减系数，值比较大，$(0.8466)^{20}=0.036$，表明冲击对上证A股指数波动造成的影响时间很长，大概需要20周的时间才能衰减到0.036。这种情况下，政策对股票市场的影响将是长期的。因此，管理层应当根据市场消化政策的能力出台相关政策，把握好调节市场的力度。

（3）GARCH-M模型残差的ARCH效应检验

为了检验GARCH-M模型是否已经消除ARCH效应检验，我们做ARCH-LM检验，滞后阶数取4，结果如图7—18所示。

ARCH Test:

F-statistic	0.728714	Probability	0.572634
Obs*R-squared	2.927187	Probability	0.570083

图7—18　ARCH-LM检验结果

可以看到ARCH-LM统计量的相伴概率为0.57，大于0.05，接受没有ARCH效应的原假设，说明残差序列不存在ARCH效应。

再观察GARCH-M模型的残差平方序列相关分析图，滞后阶数取22，如图7—19所示。

图7—19可以看到，GARCH-M模型残差平方序列的自相关系数和偏自相关系数各滞后阶数都显著为0。同时，残差平方的*Q*统计量相伴概率大于0.05。这些都说明GARCH-M模型残差序列已经不存在ARCH效应。

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.011	-0.011	0.0584	0.809
		2	0.039	0.039	0.8258	0.662
		3	-0.055	-0.054	2.3430	0.504
		4	0.039	0.037	3.1232	0.537
		5	-0.015	-0.011	3.2447	0.662
		6	-0.005	-0.011	3.2592	0.776
		7	-0.051	-0.046	4.5780	0.711
		8	-0.002	-0.005	4.5803	0.801
		9	-0.028	-0.025	4.9777	0.836
		10	-0.023	-0.028	5.2516	0.874
		11	-0.026	-0.022	5.6089	0.898
		12	0.009	0.007	5.6492	0.933
		13	-0.017	-0.017	5.7923	0.953
		14	0.054	0.050	7.3087	0.922
		15	0.045	0.049	8.3664	0.908
		16	-0.002	-0.011	8.3682	0.937
		17	-0.064	-0.065	10.517	0.881
		18	-0.001	-0.004	10.517	0.914
		19	0.012	0.012	10.587	0.937
		20	0.088	0.082	14.635	0.797
		21	-0.022	-0.013	14.892	0.828
		22	-0.043	-0.047	15.863	0.823

图 7—19 GARCH-M 模型残差平方序列的相关分析图

6. 条件异方差模型的预测

最后，我们利用加入条件标准差 $\sqrt{h_t}$ 的 GARCH-M 模型进行外推一期预测，即对 2008 年 3 月 14 日上证 A 股指数收盘价进行外推预测。方法和以前介绍的模型预测相同。

预测结果如图 7—20 所示，其中 4 395.4 为 2008 年 3 月 14 日上证 A 股指数收盘价的预测值，4 075.3～4 740.6 为加减 2 倍标准误之后的范围。GARCH 模型预测的输出结果比一般方程预测多了一项对条件方差的预测，条件方差的预测值为 0.00131。

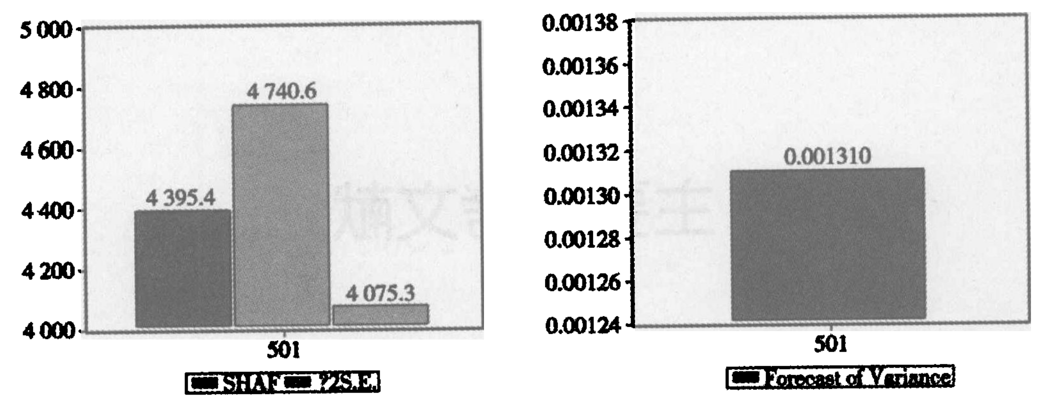


图 7—20　2008 年 3 月 14 日上证 A 股指数收盘价预测

主要参考文献

1. [美]达莫达尔·N. 古扎拉蒂：《计量经济学基础》，费剑平译，北京，中国人民大学出版社，2005。

2. [美]达莫达尔·N. 古扎拉蒂：《计量经济学精要》，张涛译，北京，机械工业出版社，2006。

3. [美]丹尼尔·L. 鲁宾费尔德、罗伯特·S. 平狄克：《计量经济模型与经济预测》，钱小军等译，北京，机械工业出版社，1999。

4. 高铁梅：《计量经济分析方法与建模——EViews 应用及实例》，北京，清华大学出版社，2006。

5. Hendry、秦朵：《动态计量经济分析》，上海，上海人民出版社，1998。

6. 李子奈、叶阿忠：《高等计量经济学》，北京，清华大学出版社，2000。

7. 陆懋祖：《高等时间序列经济计量学》，上海，上海人民出版社，1998。

8. [英]特伦斯·C. 米尔顿：《金融时间序列的经济计量学模型》，俞卓青译，北京，经济科学出版社，2002。

9. [美]威廉·H. 格林：《经济计量分析》，王明舰等译，北京，中国社会科学出版社，1998。

10. 邢天才、王玉霞：《证券投资学》，大连，东北财经大学出版社，2003。

11. 易丹辉：《数据分析与 EViews 应用》，北京，中国统计出版社，2002。

12. [美]詹姆斯·D. 汉密尔顿：《时间序列分析》，刘明志译，北京，中国社会科学出版社，1999。

13. 张雪莹、金德环：《金融计量学教程》，上海，上海财经大学出版社，2005。

14. 张晓峒：《计量经济学软件 EViews 使用指南》，天津，南开大学出版社，2004。

15. 张晓峒：《计量经济分析》，北京，经济科学出版社，2000。

16. 周爱民等：《金融计量学》，北京，经济管理出版社，2006。

书名=金融计量学实验
作者=曲春青主编
出版社=东北财经大学出版社
出版日期=2008.8
页数=142
SS号=12077498
DX号=000006603031
封面=http://cover.duxiu.com/cover/Cover.dll?iid=646562636A6A62676C6B31393831373333333
试读地址=http://book.duxiu.com/bookDetail.jsp?dxNumber=000006603031&d=3B098FE44170EC53F22D4910A12F834C